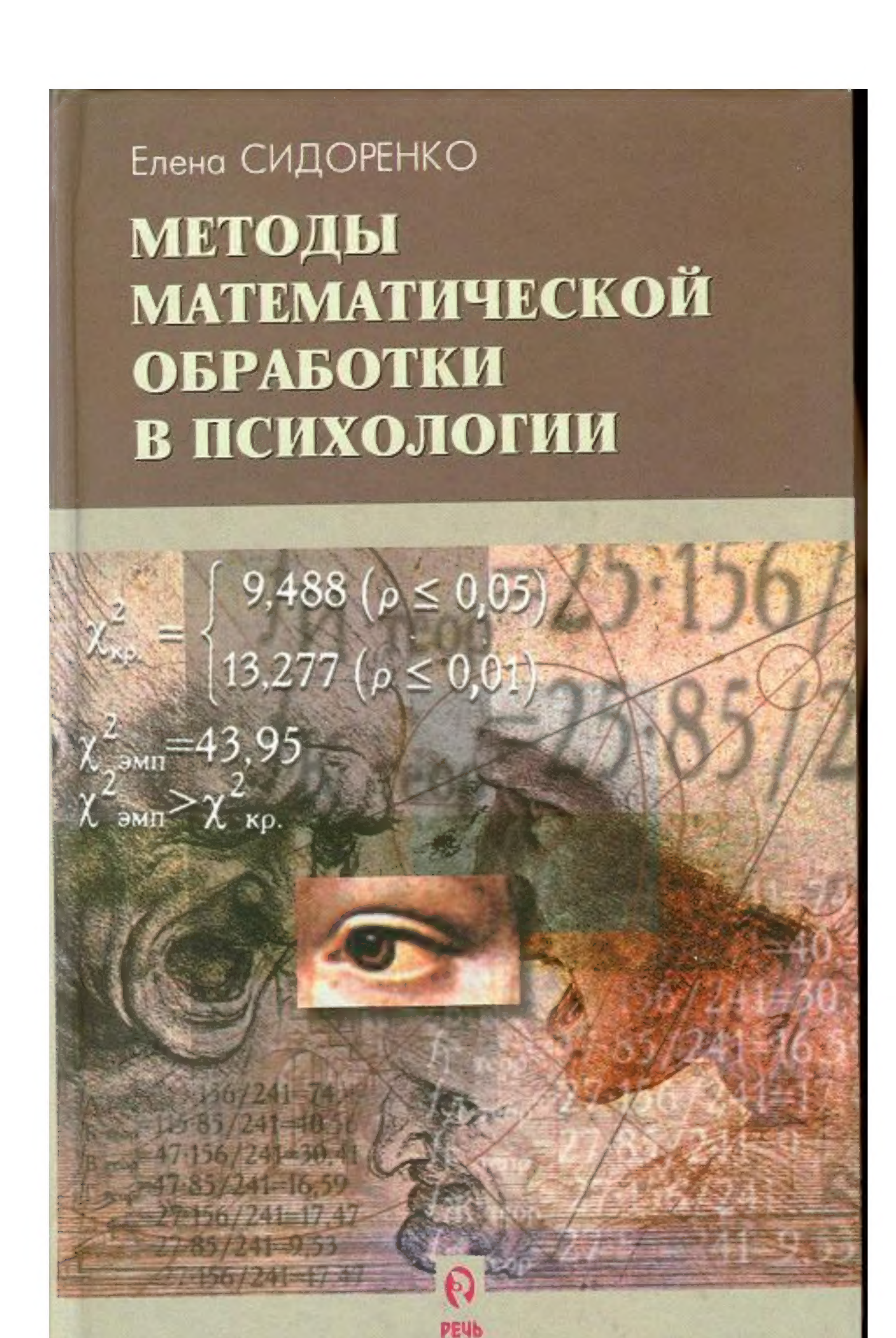

Елена СИДОРЕНКО
МЕТОДЫ МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ В ПСИХОЛОГИИ
$\chi^2_{кр.} = \begin{cases} 9,488\ (\rho \leq 0,05) \\ 13,277\ (\rho \leq 0,01) \end{cases}$
$\chi^2_{эмп}=43,95$
$\chi^2_{эмп}>\chi^2_{кр.}$
РЕЧЬ

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ АВТОРА

Принято считать, что математика - это царица наук, и любая наука становится по-настоящему наукой, только когда она начинает использовать математику. Однако многие психологи в глубине души уверены, что царица наук - отнюдь не математика, а психология. Может быть, это скорее два независимых царства, существующих как параллельные миры? Математику для доказательства своих положений совершенно не требуется привлекать психологию, а психологу можно совершать открытия, не привлекая математики. Большинство теорий личности и психотерапевтических концепций были сформулированы безо всякого обращения к математике. Примером могут служить теория психоанализа, бихевиоральная концепция, аналитическая психология К. Юнга, индивидуальная психология А. Адлера, объективная психология В.М. Бехтерева, культурно-историческая теория Л.С. Выготского, концепция отношений личности В. Н. Мясищева и многие другие теории.

Но все это было, в основном, в прошлом. Многие психологические концепции ныне подвергаются сомнению на основании того, что они не были подтверждены статистически. Стало принято использовать математические методы, как принято жениться молодому человеку, если он хочет сделать дипломатическую или политическую карьеру, и выходить замуж молодой девушке, чтобы доказать, что она может сделать это не хуже, чем все остальные. Но как не всякий молодой человек женится и не всякая девушка выходит замуж, так и не всякое психологическое исследование "венчается" с математикой.

"Брак" психологии с математикой - это брак по принуждению или недоразумению. "Глубокое внутреннее родство, общность происхождения современной физики и современной математики привели к опасному..." представлению о том, что всякое явление обязано иметь математическую модель. Это представление тем опаснее, что оно часто считается само собой разумеющимся" (А.М. Молчанов, 1978, с.4).

Психология - это невеста без приданого, у которой нет ни своих собственных единиц измерения, ни отчетливого представления о том, как заимствованные ею единицы измерения - миллиметры, секунды и градусы - соотносятся с психическими феноменами. Эти единицы измерения она взяла напрокат у физики, как отчаявшаяся бедная невеста берет взаймы подвенечное платье у более обеспеченной подруги, лишь бы царственный старец взял ее себе в младшие жены.

Между тем, "...явления, составляющие предмет гуманитарных наук, неизмеримо сложнее тех, которыми занимаются точные. Они гораздо труднее (если вообще) поддаются формализации... Вербальный способ построения исследования здесь, как это ни парадоксально, оказывается *точнее* формально-логического" (И. Грекова, 1976, с.107).

Но каковы эти вербальные способы? Какой иной язык может предложить психология вместо уже ставшего привычным языка средних, стандартных отклонений, статистически значимых различий и факторных весов? Этой задачи психология пока не решила. Уникальная специфика психологического исследования пока все еще сводится к традиционному приписыванию рангов и чисел явлениям, столь тонким, неуловимым и динамичным, что, по-видимому, к ним применима лишь принципиально иная система регистрации и оценки. Психология отчасти сама виновата в том, что ее заставляют вступать в неравный брак с математикой. Она не смогла пока еще доказать, что строится на принципиально иных основах.

Но пока психология не докажет, что может существовать независимо от математики, развод невозможен. Нам придется использовать математические методы, чтобы избавиться от необходимости объяснять, а почему мы, собственно, их не использовали? Легче использовать их, чем доказать, что в этом не было необходимости. Если же мы используем их, то целесообразно извлечь из этого максимум пользы. В любом случае, математика, несомненно, систематизирует мышление и позволяет выявить закономерности, на первый взгляд не всегда очевидные.

Ленинградская-Петербургская школа психологии, быть может, более всех других отечественных школ ориентирована на извлечение максимальной пользы из союза психологии с математикой. В 1981 году на Школе молодых ученых в Минске ленинградцы снисходительно улыбались москвичам ("Опять на одном испытуемом закономерность выстраивают!"), а москвичи - ленинградцам ("Опять своими каракатицами[1] все запутали!").

Автор этой книги принадлежит к Ленинградской психологической школе. Поэтому с первых шагов в психологии я прилежно вычисляла сигмы и подсчитывала корреляции, включала разные комбинации признаков в факторный анализ и потом ломала голову над интерпретацией факторов, обсчитывала бесконечное количество дисперсионных комплексов и др. Эти поиски продолжаются вот уже более двадцати лет. За это время я пришла к выводу, что чем проще методы математической обработки и чем ближе они к реально полученным эмпирическим данным, тем более надежными и осмысленными получаются результаты. Факторный и таксономический анализ уже слишком сложны и запутанны, чтобы каждый исследователь мог точно понимать, какие преобразования стоят за ними. Он лишь вводит свои данные в "черный ящик", а затем получает ленты машинной выдачи с факторными весами признаков, группировками испытуемых и т.д. Далее начинается интерпретация полученных факторов или классификаций, и, как любая интерпретация, она неизбежно субъективна. Но ведь субъективно судить о психических феноменах мы можем и безо всяких измерений и вычислений. Интерпретации результатов сложных обсчетов несут в себе лишь видимость научной объективности, поскольку мы по-прежнему субъективно интерпретируем, но уже не реальные результаты наблюдений, а результаты их математической обработки. По этой причине факторный, дискриминантный, кластерный, таксономический виды анализа не рассматриваются мною в этой книге.

Принцип отбора методов в данном руководстве - простота и практичность. Большинство методов построены на понятных для исследователя преобразованиях. Некоторые из них ранее редко использовались или не использовались совсем - например, критерий тенденций S Джонкира и L Пейджа. Они могут рассматриваться как эффективная замена метода линейной корреляции.

Большинство рассматриваемых методов являются непараметрическими, или "свободными от распределения", что значительно расширяет их возможности по сравнению с традиционными параметрическими методами, например t - критерием Стьюдента и методом линейной корреляции Пирсона. Некоторые из предлагаемых методов могут быть применены по отношению к любым данным, имеющим хоть какое-то числовое выражение. Принцип каждого метода иллюстрируется графически, с тем, чтобы всякий раз исследователь отчетливо осознавал, какого рода преобразования он совершает.

Все методы рассматриваются на примерах, полученных в реальных психологических исследованиях. К Главам 2-5 прилагаются задачи для самостоятельной работы, решение которых подробно рассматривается в Главе 9.

Все представленные экспериментальные результаты могут использоваться для научных сопоставлений, так как это реальные научные данные, полученные мною в собственных исследованиях, в совместных исследованиях с моими коллегами или моими учениками.

Применение реальных данных позволяет избежать тех несообразностей, которые часто возникают при рассмотрении искусственно придуманных задач. Принцип реальности позволяет по-настоящему почувствовать подводные камни и тонкости в использовании статистических методов и интерпретации полученных результатов.

Выражаю глубокую признательность людям, без встречи с которыми эта книга не была бы написана. Прежде всего - моим учителям в области математики и математической статистики, Инне Леонидовне Улитиной и профессору Геннадию

[1] "Каракатица" - ироническое обозначение корреляционной плеяды.

Владимировичу Суходольскому, благодаря которым использование математики стало для меня скорее удовольствием, чем неприятной обязанностью.

Погрузиться в таинственный мир психологического эксперимента и почувствовать "вкус" к поиску статистических закономерностей мне помогли в юности мои старшие коллеги по Лаборатории антропологии и дифференциальной психологии имени академика Б.Г. Ананьева: Мария Дмитриевна Дворяшина, Борис Степанович Одерышев, Владимир Константинович Горбачевский, Людмила Николаевна Кулешова, Иосиф Маркович Палей, Галина Ивановна Акинщикова, Елена Федоровна Рыбалко, Нина Альбертовна Грищенко-Розе, Лариса Арсеньевна Го-ловей, Николай Николаевич Обозов, Нина Михайловна Владимирова, Ольга Михайловна Анисимова, позже, уже в Лаборатории экспериментальной и прикладной психологии - Капитолина Дмитриевна Шафран-ская.

Все эти люди были влюблены в психологию. Увлеченно и страстно они старались проникнуть в суть того, что проявляется на поверхности человеческих действий и реакций. Воспоминания о совместных поисках и находках неизменно вдохновляли меня при написании этой книги.

Я глубоко благодарна своему научному руководителю по аспирантуре - декану факультета психологии Санкт-Петербургского университета профессору Альберту Александровичу Крылову - за способность передать мне ощущение гармонии эмпирического материала и за мудрое требование переводить абстрактные математические результаты на язык графических образов, возвращающих к исследуемой реальности.

В разные годы мне очень помогли своими математическими советами психологи: Аркадий Ильич Нафтульев и Наталия Марковна Лебедева, - и математики: Владимир Филиппович Федоров, Михаил Александрович Скороденок, Ярослав Александрович Бедров, Вячеслав Леонидович Кузнецов, Елена Андреевна Вершинина и математический редактор этого руководства Александр Борисович Алексеев, чьи консультации и поддержка были необходимы, как воздух, при подготовке книги.

Выражаю свою признательность руководителю Вычислительного Центра факультета Михаилу Михайловичу Зиберту и сотрудникам центра - Эльвире Аркадьевне Яковлевой, Татьяне Ивановне Гусевой, Григорию Петровичу Савченко за неоценимую помощь в подготовке программ и обработке моих материалов на протяжение многих лет.

В моем сердце жива благодарность и к тем коллегам, кого уже нет с нами - Надежде Петровне Чумаковой, Виктору Ивановичу Бутову, Белле Ефимовне Шустер. Их дружеская поддержка и профессиональная помощь были неоценимы.

Я отдаю глубокую дань памяти Евгению Сергеевичу Кузьмину, возглавлявшему кафедру социальной психологии Санкт-Петербургского университета в 1966-1988 годах и разработавшему целостную концепцию теоретической и практической подготовки социальных психологов, в программу которой вошел и лекционно-практический курс "Методы математической обработки в психологических исследованиях". Я благодарна ему за включение меня в свой замечательный коллектив, доброе уважительное отношение ко мне и веру в мои профессиональные возможности.

И, наконец, последнее — по списку, но не по значению. Я глубоко благодарна нынешнему заведующему кафедрой социальной психологии — профессору Анатолию Леонидовичу Свенцицкому — за открытость новым идеям и поддержание на кафедре атмосферы свободного поиска, высоких интеллектуальных требований и дружеской поддержки, окрашенной юмором и мягкой иронией. Именно такая среда вдохновляет к творчеству.

Как читать эту книгу и как ею пользоваться

Начинающим лучше начать чтение с Главы 1 , затем выбрать, на основании алгоритмов 1 и 2, какой метод им лучше использовать, *разобраться в примере.* Затем стоит внимательно прочитать весь параграф, относящийся к данному методу, и

попробовать самостоятельно решить прилагаемые задачи. После этого можно смело начать решение собственной задачи или... переключиться на другой метод, если Вы убедились, что этот Вам не подходит.

Знатокам можно сразу обращаться к методам, которые кажутся им подходящими для их задачи. Они могут *использовать алгоритм* применения избранного метода или опираться на пример, как нечто более наглядное. Для интерпретации результатов им, возможно, понадобится познакомиться с разделом "Графическое представление критерия". Не исключено, что анализ задач, предлагаемых в руководстве, поможет им увидеть новые грани в использовании знакомого метода.

Владельцам компьютерных программ подсчета статистических критериев может оказаться необходимым *познакомиться* с *идеологией* избранного ими метода в разделах "Описание", "Гипотезы", "Ограничения" и "Графическое представление критерия" — ведь компьютер не объясняет, каковы способы интерпретации полученных числовых значений.

Стремящимся к быстроте лучше сразу обращаться к п. 5.2 о критерии φ* (угловое преобразование Фишера). Этот метод поможет решить почти любую задачу.

Стремящимся к основательности можно прочитать, помимо прочего, также и те разделы текста, которые набраны мелким шрифтом.

Желаю успеха!

Елена Сидоренко

ГЛАВА 1
ОСНОВНЫЕ ПОНЯТИЯ, ИСПОЛЬЗУЕМЫЕ В МАТЕМАТИЧЕСКОЙ ОБРАБОТКЕ ПСИХОЛОГИЧЕСКИХ ДАННЫХ

1.1. Признаки и переменные

Признаки и переменные - это измеряемые психологические явления. Такими явлениями могут быть время решения задачи, количество допущенных ошибок, уровень тревожности, показатель интеллектуальной лабильности, интенсивность агрессивных реакций, угол поворота корпуса в беседе, показатель социометрического статуса и множество других переменных.

Понятия признака и переменной могут использоваться как взаимозаменяемые. Они являются наиболее общими. Иногда вместо них используются понятия показателя или уровня, например, уровень настойчивости, показатель вербального интеллекта и др. Понятия показателя и уровня указывают на то, что признак может быть измерен количественно, так как к ним применимы определения "высокий" или "низкий", например, высокий уровень интеллекта, низкие показатели тревожности и др.

Психологические переменные являются случайными величинами, поскольку заранее неизвестно, какое именно значение они примут.

Математическая обработка - это оперирование со значениями признака, полученными у испытуемых в психологическом исследовании. Такие индивидуальные результаты называют также "наблюдениями", "наблюдаемыми значениями", "вариантами", "датами", "индивидуальными показателями" и др. В психологии чаще всего используются термины "наблюдение" или "наблюдаемое значение".

Значения признака определяются при помощи специальных шкал измерения.

1.2. Шкалы измерения

Измерение - это приписывание числовых форм объектам или событиям в соответствии с определенными правилами (Стивене С, 1960, с.60). С.Стивенсом предложена классификация из 4 типов шкал измерения:

1) номинативная, или номинальная, или шкала наименований;
2) порядковая, или ординальная, шкала;
3) интервальная, или шкала равных интервалов;
4) шкала равных отношений.

Номинативная шкала - это шкала, классифицирующая по названию: *nomen* (лат.) - имя, название. Название же не измеряется количественно, оно лишь позволяет отличить один объект от другого или одного субъекта от другого. Номинативная шкала - это способ классификации объектов или субъектов, распределения их по ячейкам классификации.

Простейший случай номинативной шкалы - дихотомическая шкала, состоящая всего лишь из двух ячеек, например: "имеет братьев и сестер - единственный ребенок в семье"; "иностранец - соотечественник"; "проголосовал "за" - проголосовал "против"" и т.п.

Признак, который измеряется по дихотомической шкале наименований, называется альтернативным. Он может принимать всего два значения. При этом исследователь зачастую заинтересован в одном из них, и тогда он говорит, что признак “проявился”, если тот принял интересующее его значение, и что признак “не проявился” , если он принял противоположное значение. Например: "Признак леворукости проявился у 8 испытуемых из 20". В принципе номинативная шкала может состоять из ячеек "признак проявился - признак не проявился .

Более сложный вариант номинативной шкалы - классификация из трех и более ячеек, например: "экстрапунитивные - интрапунитивные -импунитивные реакции" или "выбор кандидатуры А - кандидатуры Б -кандидатуры В - кандидатуры Г" или "старший - средний - младший -единственный ребенок в семье" и др.

Расклассифицировав все объекты, реакции или всех испытуемых по ячейкам классификации, мы получаем возможность от наименований перейти к числам, подсчитав количество наблюдений в каждой из ячеек.

Как уже указывалось, наблюдение - это одна зарегистрированная реакция, один совершенный выбор, одно осуществленное действие или результат одного испытуемого.

Допустим, мы определим, что кандидатуру А выбрали 7 испытуемых, кандидатуру Б - 11, кандидатуру В - 28, а кандидатуру Г -всего 1. Теперь мы можем оперировать этими числами, представляющими собой частоты встречаемости разных наименований, то есть частоты принятия признаком "выбор" каждого из 4 возможных значений. Далее мы можем сопоставить полученное распределение частот с равномерным или каким-то иным распределением.

Таким образом, номинативная шкала позволяет нам подсчитывать частоты встречаемости разных "наименований", или значений признака, и затем работать с этими частотами с помощью математических методов.

Единица измерения, которой мы при этом оперируем - количество наблюдений (испытуемых, реакций, выборов и т. п.), или частота. Точнее, единица измерения - это одно наблюдение. Такие данные могут быть обработаны с помощью метода χ^2, биномиального критерия m и углового преобразования Фишера φ*.

Порядковая шкала - это шкала, классифицирующая по принципу "больше - меньше". Если в шкале наименований было безразлично, в каком порядке мы расположим классификационные ячейки, то в порядковой шкале они образуют последовательность от ячейки "самое малое значение" к ячейке "самое большое значение" (или наоборот). Ячейки теперь уместнее называть классами, поскольку по отношению к классам употребимы определения "низкий", "средний" и "высокий" класс, или 1-й, 2-й, 3-й класс, и т.д.

В порядковой шкале должно быть не менее трех классов, например "положительная реакция - нейтральная реакция - отрицательная реакция" или "подходит для занятия вакантной должности - подходит с оговорками - не подходит" и т. п.

В порядковой шкале мы не знаем истинного расстояния между классами, а знаем лишь, что они образуют последовательность. Например, классы "подходит для занятия вакантной должности" и "подходит с оговорками" могут быть реально ближе друг к другу, чем класс "подходит с оговорками" к классу "не подходит".

От классов легко перейти к числам, если мы условимся считать, что низший класс получает ранг 1, средний класс - ранг 2, а высший класс - ранг 3, или наоборот. Чем

больше классов в шкале, тем больше у нас возможностей для математической обработки полученных данных и проверки статистических гипотез.

Например, мы можем оценить различия между двумя выборками испытуемых по преобладанию у них более высоких или более низких рангов или подсчитать коэффициент ранговой корреляции между двумя переменными, измеренными в порядковой шкале, допустим, между оценками профессиональной компетентности руководителя, данными ему разными экспертами.

Все психологические методы, использующие ранжирование, построены на применении шкалы порядка. Если испытуемому предлагается упорядочить 18 ценностей по степени их значимости для него, проранжировать список личностных качеств социального работника или 10 претендентов на эту должность по степени их профессиональной пригодности, то во всех этих случаях испытуемый совершает так называемое принудительное ранжирование, при котором количество рангов соответствует количеству ранжируемых субъектов или объектов (ценностей, качеств и т.п.).

Независимо от того, приписываем ли мы каждому качеству или испытуемому один из 3-4 рангов или совершаем процедуру принудительного ранжирования, мы получаем в обоих случаях ряды значений, измеренные по порядковой шкале. Правда, если у нас всего 3 возможных класса и, следовательно, 3 ранга, и при этом, скажем, 20 ранжируемых испытуемых, то некоторые из них неизбежно получат одинаковые ранги. Все многообразие жизни не может уместиться в 3 градации, поэтому в один и тот же класс могут попасть люди, достаточно серьезно различающиеся между собой. С другой стороны, принудительное ранжирование, то есть образование последовательности из многих испытуемых, может искусственно преувеличивать различия между людьми. Кроме того, данные, полученные в разных группах, могут оказаться несопоставимыми, так как группы могут изначально различаться по уровню развития исследуемого качества, и испытуемый, получивший в одной группе высший ранг, в другой получил бы всего лишь средний, и т.п.

Выход из положения может быть найден, если задавать достаточно дробную классификационную систему, скажем, из 10 классов, или градаций, признака. В сущности, подавляющее большинство психологических методик, использующих экспертную оценку, построено на измерении одним и тем же "аршином" из 10, 20 или даже 100 градаций разных испытуемых в разных выборках.

Итак, единица измерения в шкале порядка - расстояние в 1 класс или в 1 ранг, при этом расстояние между классами и рангами может быть разным (оно нам неизвестно). К данным, полученным по порядковой шкале, применимы все описанные в данной книге критерии и методы.

Интервальная шкала - это шкала, классифицирующая по принципу "больше на определенное количество единиц - меньше на определенное количество единиц". Каждое из возможных значений признака отстоит от другого на равном расстоянии.

Можно предположить, что если мы измеряем время решения задачи в секундах, то это уже явно шкала интервалов. Однако на самом деле это не так, поскольку психологически различие в 20 секунд между испытуемым А и Б может быть отнюдь не равно различию в 20 секунд между испытуемыми Б и Г, если испытуемый А решил задачу за 2 секунды, Б - за 22, В - за 222, а Г - за 242.

Аналогичным образом, каждая секунда после истечения полутора минут в опыте с измерением мышечного волевого усилия на динамометре с подвижной стрелкой, по "цене", может быть, равна 10 или даже более секундам в первые полминуты опыта. "Одна секунда за год идет" - так сформулировал это однажды один испытуемый.

Попытки измерять психологические явления в физических единицах - волю в секундах, способности в сантиметрах, а ощущение собственной недостаточности - в миллиметрах и т. п., конечно, понятны, ведь все-таки это измерения в единицах "объективно" существующего времени и пространства. Однако ни один опытный

исследователь при этом не обольщает себя мыслью, что он совершает измерения по психологической интервальной шкале. Эти измерения принадлежат по-прежнему к шкале порядка, нравится нам это или нет (Стивене С, 1960, с.56; Паповян С.С., 1983, с.63; Михеев В.И.: 1986, с.28).

Мы можем с определенной долей уверенности утверждать лишь, что испытуемый А решил задачу быстрее Б, Б быстрее В, а В быстрее Г.

Аналогичным образом, значения, полученные испытуемыми в баллах по любой нестандартизованной методике, оказываются измеренными лишь по шкале порядка. На самом деле равноинтервальными можно считать лишь шкалы в единицах стандартного отклонения и про-центильные шкалы, и то лишь при условии, что распределение значений в стандартизующей выборке было нормальным (Бурлачук Л. Ф., Морозов С. М., 1989, с. 163, с. 101).

Принцип построения большинства интервальных шкал построен на известном правиле "трех сигм": примерно 97,7-97,8% всех значений признака при нормальном его распределении укладываются в диапазоне $M \pm 3\sigma$[2] Можно построить шкалу в единицах долей стандартного отклонения, которая будет охватывать весь возможный диапазон изменения признака, если крайний слева и крайний справа интервалы оставить открытыми.

Р.Б. Кеттелл предложил, например, шкалу стенов - "стандартной десятки". Среднее арифметическое значение в "сырых" баллах принимается за точку отсчета. Вправо и влево отмеряются интервалы, равные 1/2 стандартного отклонения. На Рис. 1.2 представлена схема вычисления стандартных оценок и перевода "сырых" баллов в стены по шкале N 16-факторного личностного опросника Р. Б. Кеттелла.

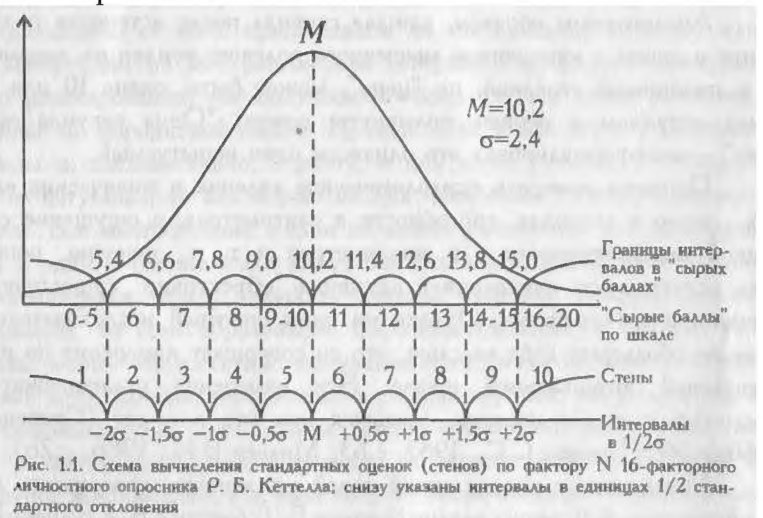


Рис. 1.1. Схема вычисления стандартных оценок (стенов) по фактору N 16-факторного личностного опросника Р. Б. Кеттелла; снизу указаны интервалы в единицах 1/2 стандартного отклонения

Справа от среднего значения будут располагаться интервалы, равные 6, 7, 8, 9 и 10 стенам, причем последний из этих интервалов открыт. Слева от среднего значения будут располагаться интервалы, равные 5, 4, 3, 2 и 1 стенам, и крайний интервал также открыт. Теперь мы поднимаемся вверх, к оси "сырых баллов", и размечаем границы интервалов в единицах "сырых" баллов. Поскольку M=10,2; σ=2,4, вправо мы откладываем 1/2σ, т.е. 1,2 "сырых" балла. Таким образом, граница интервала составит: (10,2 + 1,2) = 11,4 "сырых" балла. Итак, границы интервала, соответствующего 6 стенам, будут простираться от 10,2 до 11,4 баллов. В сущности, в него попадает только одно "сырое" значение - 11 баллов. Влево от средней мы откладываем 1/2 σ и получаем границу интервала: 10,2-1,2=9. Таким образом, границы интервала, соответствующие 9 стенам, простираются от 9 до 10,2. В этот интервал попадают уже два "сырых" значения - 9 и 10. Если испытуемый получил 9 "сырых" баллов, ему начисляется теперь 5 стенов; если он получил 11 "сырых" баллов - 6 стенов, и т. д.

Мы видим, что в шкале стенов иногда за разное количество "сырых" баллов будет начисляться одинаковое количество стенов. Например, за 16, 17, 18, 19 и 20 баллов будет начисляться 10 стенов, а за 14 и 15 - 9 стенов и т. д.

В принципе,шкалу стенов можно построить по любым данным, измеренным по крайней мере в

[2] Определения и формулы расчета M и СТ даны в параграфе "Распределение признака. Параметры распределения".

порядковой шкале, при объеме выборки п>200 и нормальном распределении признака[3].

Другой способ построения равноинтервальной шкалы - группировка интервалов по принципу равенства накопленных частот. При нормальном распределении признака в окрестности среднего значения группируется большая часть всех наблюдений, поэтому в этой области среднего значения интервалы оказываются меньше, уже, а по мере удаления от центра распределения они увеличиваются, (см. Рис. 1.2). Следовательно, такая процентильная шкала является равноинтервальной только относительно накопленной частоты (Мельников В.М., Ямпольский Л.Т., 1985, с. 194).

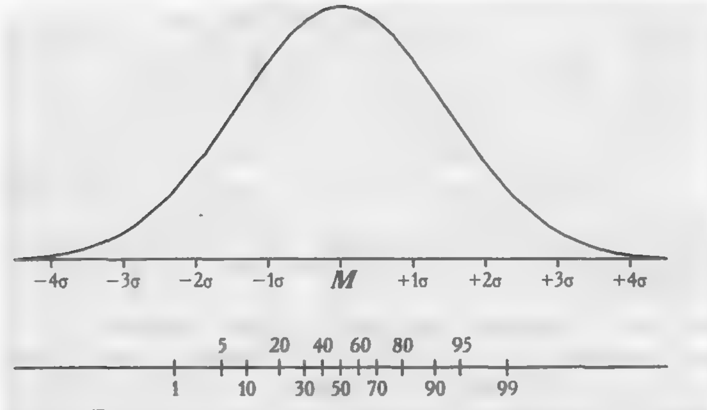


Рис. 1.2. Процентильная шкала; сверху для сравнения указаны интервалы в единицах стандартного отклонения

Построение шкал равных интервалов по данным, полученным по шкале порядка, напоминает трюк с веревочной лестницей, на который ссылался С. Стивене. Мы сначала поднимаемся по лестнице, которая ни на чем не закреплена, и добираемся до лестницы, которая закреплена. Однако каким путем мы оказались на ней? Измерили некую психологическую переменную по шкале порядка, подсчитали средние и стандартные отклонения, а затем получили, наконец, интервальную шкалу. "Такому нелегальному использованию статистики может быть дано известное прагматическое оправдание; во многих случаях оно приводит к плодотворным результатам" (Стивене С, 1960, с. 56).

Многие исследователи не проверяют степень совпадения полученного ими эмпирического распределения с нормальным распределением, и тем более не переводят получаемые значения в единицы долей стандартного отклонения или процентили, предпочитая пользоваться "сырыми" данными. "Сырые" же данные часто дают скошенное, срезанное по краям или двухвершинное распределение. На Рис. 1.3 представлено распределение показателя мышечного волевого усилия на выборке из 102 испытуемых. Распределение с удовлетворительной точностью можно считать нормальным (χ^2=12,7, при v=9, M=89,75, σ= 25,1).

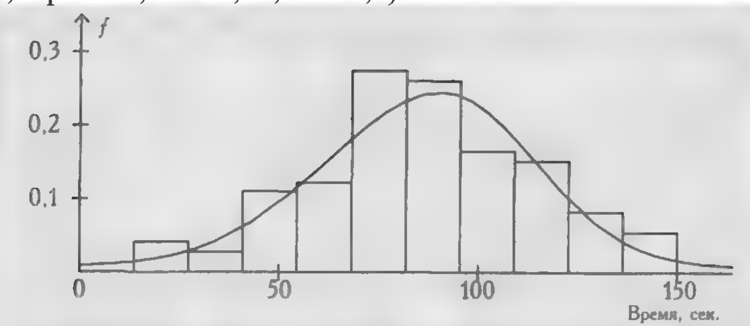


Рис. 1.3. Гистограмма и плавная кривая распределения показателя мышечного волевого усилия (n=102)

На Рис. 1.4 представлено распределение показателя самооценки по шкале методики Дж. Менестера - Р.Корзини "Уровень успеха, которого я должен был достичь уже сейчас" (n=356). Распределение значимо отличается от нормального (χ^2=58,8, при v=7; p<0,01; M=80,64; σ =16,86).

[3] О нормальном распределении см. Пояснения в п. 1.3.

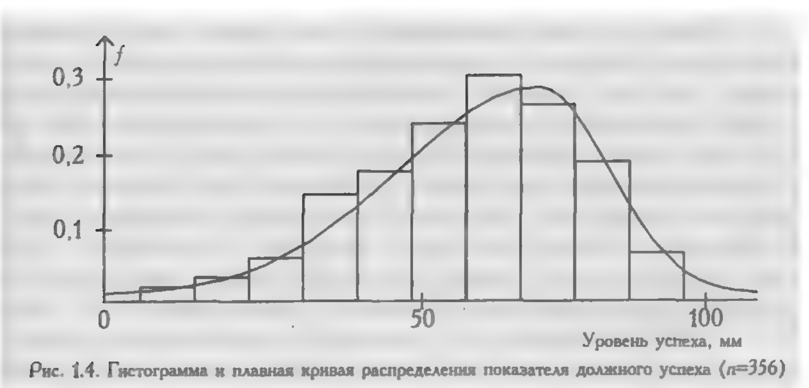


Рис. 1.4. Гистограмма и плавная кривая распределения показателя должного успеха (n=356)

С такими "ненормальными" распределениями приходится встречаться очень часто, чаще, может быть, чем с классическими нормальными. И дело здесь не в каком-то изъяне, а в самой специфике психологических признаков. По некоторым методикам от 10 до 20% испытуемых получают оценку "ноль" - например, в их рассказах не встречается ни одной словесной формулировки, которая отражала бы мотив "надежда на успех" или "боязнь неудачи" (методика Хекхаузена). То, что испытуемый получил оценку "ноль", нормально, но распределение таких оценок не может быть нормальным, как бы мы ни увеличивали объем выборки (см. п. 5.3).

Методы статистической обработки, предлагаемые в настоящем руководстве, в большинстве своем не требуют проверки совпадения полученного эмпирического распределения с нормальным. Они построены на подсчете частот и ранжировании. Проверка необходима только в случае применения дисперсионного анализа. Именно поэтому соответствующая глава сопровождается описанием процедуры подсчета необходимых критериев.

Во всех остальных случаях нет необходимости проверять степень совпадения полученного эмпирического распределения с нормальным, и тем более стремиться преобразовать порядковую шкалу в равноинтервальную. В каких бы единицах ни были измерены переменные - в секундах, миллиметрах, градусах, количестве выборов и т. п. - все эти данные могут быть обработаны с помощь непараметрических критериев[4], составляющих основу данного руководства.

Шкала равных отношений - это шкала, классифицирующая объекты или субъектов пропорционально степени выраженности измеряемого свойства. В шкалах отношений классы обозначаются числами, которые пропорциональны друг другу: 2 так относится к 4, как 4 к 8. Это предполагает наличие абсолютной нулевой точки отсчета. В физике абсолютная нулевая точка отсчета встречается при измерении длин отрезков или физических объектов и при измерении температуры по шкале Кельвина с абсолютным нулем температур. Считается, что в психологии примерами шкал равных отношений являются шкалы порогов абсолютной чувствительности (Стивене С, 1960; Гайда В. К., Захаров В. П., 1982). Возможности человеческой психики столь велики, что трудно представить себе абсолютный нуль в какой-либо измеряемой психологической переменной. Абсолютная глупость и абсолютная честность - понятия скорее житейской психологии.

То же относится и к установлению равных отношений: только метафора обыденной речи допускает, чтобы Иванов был в 2 раза (3, 100, 1000) умнее Петрова или наоборот.

Абсолютный нуль, правда, может иметь место при подсчете количества объектов или субъектов. Например, при выборе одной из 3 альтернатив испытуемые не выбрали альтернативу А ни одного раза, альтернативу Б - 14 раз и альтернативу В - 28 раз. В этом случае мы можем утверждать, что альтернативу В выбирают в два раза чаще, чем альтернативу Б. Однако при этом измерено не психологическое свойство человека, а соотношение выборов у 42 человек.

[4] Определение и описание непараметрических критериев дано ниже в данной главе.

По отношению к показателям частот возможно применять все арифметические операции: сложение, вычитание, деление и умножение. Единица измерения в этой шкале отношений - 1 наблюдение, 1 выбор, 1 реакция и т. п. Мы вернулись к тому, с чего начали: к универсальной шкале измерения в частотах встречаемости того или иного значения признака и к единице измерения, которая представляет собой 1 наблюдение. Расклассифицировав испытуемых по ячейкам номинативной шкалы, мы можем применить потом высшую шкалу измерения - шкалу отношений между частотами.

1.3. Распределение признака. Параметры распределения

Распределением признака называется закономерность встречаемости разных его значений (Плохинский Н.А., 1970, с. 12).

В психологических исследованиях чаще всего ссылаются на нормальное распределение.

Нормальное распределение характеризуется тем, что крайние значения признака в нем встречаются достаточно редко, а значения, близкие к средней величине - достаточно часто. Нормальным такое распределение называется потому, что оно очень часто встречалось в естественнонаучных исследованиях и казалось "нормой" всякого массового случайного проявления признаков. Это распределение следует закону, открытому тремя учеными в разное время: Муавром в 1733 г. в Англии, Гауссом в 1809 г. в Германии и Лапласом в 1812 г. во Франции (Плохинский Н.А., 1970, с.17). График нормального распределения представляет собой привычную глазу психолога-исследователя так называемую колоколообразную кривую (см. напр., Рис. 1.1, 1.2).

Параметры распределения - это его числовые характеристики, указывающие, где "в среднем" располагаются значения признака, насколько эти значения изменчивы и наблюдается ли преимущественное появление определенных значений признака. Наиболее практически важными параметрами являются математическое ожидание, дисперсия, показатели асимметрии и эксцесса.

В реальных психологических исследованиях мы оперируем не параметрами, а их приближенными значениями, так называемыми оценками параметров. Это объясняется ограниченностью обследованных выборок. Чем больше выборка, тем ближе может быть оценка параметра к его истинному значению. В дальнейшем, говоря о параметрах, мы будем иметь в виду их оценки.

Среднее арифметическое (оценка математического ожидания) вычисляется по формуле:

$$\bar{x} = M = \frac{\sum x_i}{n}$$

где x_i - каждое наблюдаемое значение признака;

i - индекс, указывающий на порядковый номер данного значения признака;

n - количество наблюдений;

$\sum$- знак суммирования.

Оценка дисперсии определяется по формуле:

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

где x_i - каждое наблюдаемое значение признака;

$\bar{x}$ - среднее арифметическое значение признака;

n - количество наблюдений.

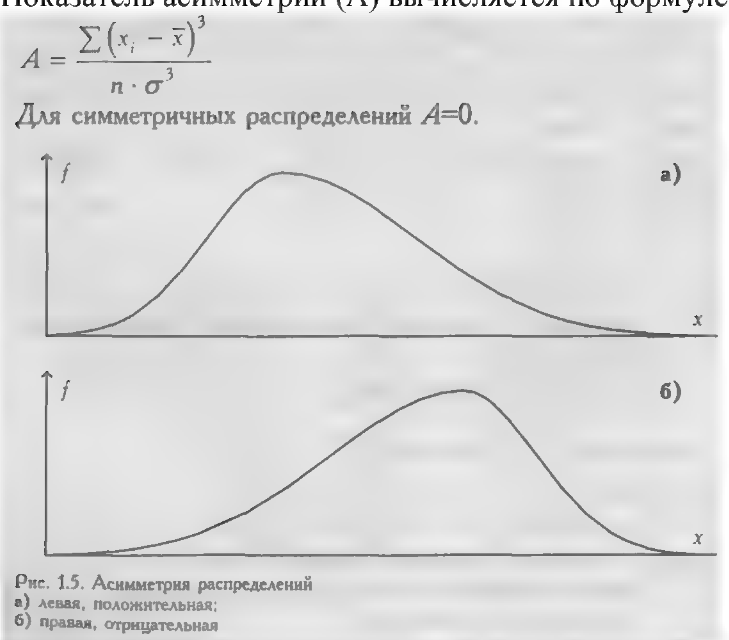
Величина, представляющая собой квадратный корень из несмещенной оценки дисперсии (S), называется стандартным отклонением или средним квадратическим отклонением. Для большинства исследователей привычно обозначать эту величину греческой буквой σ (сигма), а не S. На самом деле, σ - это стандартное отклонение в генеральной совокупности, а S - несмещенная оценка этого параметра в исследованной выборке. Но, поскольку S - лучшая оценка σ (Fisher R.A., 1938), эту оценку стали часто

обозначать уже не как S, а как σ:

$$\sigma = \sqrt{\frac{\sum\left(x_i - \bar{x}\right)^2}{n-1}}$$

В тех случаях, когда какие-нибудь причины благоприятствуют более частому появлению значений, которые выше или, наоборот, ниже среднего, образуются асимметричные распределения. При левосторонней, или положительной, асимметрии в распределении чаще встречаются более низкие значения признака, а при правосторонней, или отрицательной - более высокие (см. Рис. 1.5).

Показатель асимметрии (А) вычисляется по формуле:



$$A = \frac{\sum\left(x_i - \bar{x}\right)^3}{n \cdot \sigma^3}$$

Для симметричных распределений A=0.

Рис. 1.5. Асимметрия распределений
а) левая, положительная;
б) правая, отрицательная

В тех случаях, когда какие-либо причины способствуют преимущественному появлению средних или близких к средним значений, образуется распределение с положительным эксцессом. Если же в распределении преобладают крайние значения, причем одновременно и более низкие, и более высокие, то такое распределение характеризуется отрицательным эксцессом и в центре распределения может образоваться впадина, превращающая его в двувершинное (см. Рис. 1.6).

Показатель эксцесса (Е) определяется по формуле:

$$E = \frac{\sum(x_i - \bar{x})^4}{n \cdot \sigma^4} - 3$$

Рис. 1.6. Эксцесс: а) положительный; б) отрицательный

В распределениях с нормальной выпуклостью E=0.

Параметры распределения оказывается возможным определить только по отношению к данным, представленным по крайней мере в интервальной шкале. Как мы убедились ранее, физические шкалы длин, времени, углов являются интервальными шкалами, и поэтому к ним применимы способы расчета оценок параметров, по крайней мере, с формальной точки зрения. Параметры распределения не учитывают истинной психологической неравномерности секунд, миллиметров и других физических единиц измерения.

На практике психолог-исследователь может рассчитывать параметры любого распределения, если единицы, которые он использовал при измерении, признаются разумными в научном сообществе.

1.4. Статистические гипотезы

Формулирование гипотез систематизирует предположения исследователя и представляет их в четком и лаконичном виде. Благодаря гипотезам исследователь не теряет путеводной нити в процессе расчетов и ему легко понять после их окончания, что, собственно, он обнаружил.

Статистические гипотезы подразделяются на нулевые и альтернативные, направленные и ненаправленные.

Нулевая гипотеза -	это гипотеза об отсутствии различий. Она обозначается как H_0 и называется нулевой потому, что содержит число 0: X_1-X_2=0, где X_1, X_2 - сопоставляемые значения признаков. Нулевая гипотеза - это то, что мы хотим опровергнуть, если перед нами стоит задача доказать значимость различий.
Альтернативная гипотеза -	это гипотеза о значимости различий. Она обозначается как H_1. Альтернативная гипотеза - это то, что мы хотим доказать, поэтому иногда ее называют *экспериментальной* гипотезой.

Бывают задачи, когда мы хотим доказать как раз незначимость различий, то есть подтвердить нулевую гипотезу. Например, если нам нужно убедиться, что разные испытуемые получают хотя и различные, но уравновешенные по трудности задания, или что экспериментальная и контрольная выборки не различаются между собой по каким-то значимым характеристикам. Однако чаще нам все-таки требуется доказать *значимость различий,* ибо они более информативны для нас в поиске нового. Нулевая и альтернативная гипотезы могут быть направленными и ненаправленными.

Направленные гипотезы

H_0: X_1 не превышает X_2

H_1: X_1 превышает X_2

Ненаправленные гипотезы

H_0: X_1 не отличается от X_2

H_1: X_1 отличается от X_2

Если вы заметили, что в одной из групп индивидуальные значения испытуемых по какому-либо признаку, например по социальной смелости, выше, а в другой ниже, то для проверки значимости этих различий нам необходимо сформулировать направленные гипотезы.

Если мы хотим доказать, что в группе А под влиянием каких-то экспериментальных воздействий произошли более выраженные изменения, чем в группе Б, то нам тоже

необходимо сформулировать направленные гипотезы.

Если же мы хотим доказать, что различаются формы распределения признака в группе А и Б, то формулируются ненаправленные гипотезы.

При описании каждого критерия в руководстве даны формулировки гипотез, которые он помогает нам проверить.

Построим схему - классификацию статистических гипотез.



Проверка гипотез осуществляется с помощью критериев статистической оценки различий.

1.5. Статистические критерии

Статистический критерий - это решающее правило, обеспечивающее надежное поведение, то есть принятие истинной и отклонение ложной гипотезы с высокой вероятностью (Суходольский Г.В., 1972, с. 291).

Статистические критерии обозначают также метод расчета определенного числа и само это число.

Когда мы говорим, что достоверность различий определялась по критерию χ^2, то имеем в виду, что использовали метод χ^2 - для расчета определенного числа.

Когда мы говорим, далее, что χ^2=12,676, то имеем в виду определенное число, рассчитанное по методу χ^2. Это число обозначается как эмпирическое значение критерия.

По соотношению эмпирического и критического значений критерия мы можем судить о том, подтверждается ли или опровергается нулевая гипотеза. Например, если $\chi^2_{эмп} > \chi^2_{кр}$, H_0 отвергается.

В большинстве случаев для того, чтобы мы признали различия значимыми, необходимо, чтобы эмпирическое значение критерия превышало критическое, хотя есть критерии (например, критерий Манна-Уитни или критерий знаков), в которых мы должны придерживаться противоположного правила.

Эти правила оговариваются в описании каждого из представленных в руководстве критериев.

В некоторых случаях расчетная формула критерия включает в себя количество наблюдений в исследуемой выборке, обозначаемое как п. В этом случае эмпирическое значение критерия одновременно является тестом для проверки статистических гипотез. По специальной таблице мы определяем, какому уровню статистической значимости различий соответствует данная эмпирическая величина. Примером такого критерия является критерий φ*, вычисляемый на основе углового преобразования Шишера.

В большинстве случаев, однако, одно и то же эмпирическое значение критерия может оказаться значимым или незначимым в зависимости от количества наблюдений в исследуемой выборке (n) или от так называемого количества степеней свободы, которое обозначается как *v* или как *df.*

Число степеней свободы *v* равно числу классов вариационного ряда минус число условий, при которых он был сформирован (Ивантер Э.В., Коросов А.В., 1992, с. 56). К числу таких условий относятся объем выборки (n), средние и дисперсии.

Если мы расклассифицировали наблюдения по классам какой-либо номинативной шкалы и подсчитали количество наблюдений в каждой ячейке классификации, то мы получаем так называемый частотный вариационный ряд. Единственное условие, которое соблюдается при его формировании - объем выборки *n.* Допустим, у нас 3 класса: "Умеет работать на компьютере - умеет выполнять лишь определенные операции - не умеет работать на компьютере". Выборка состоит из 50 человек. Если в первый класс отнесены

20 испытуемых, во второй - тоже 20, то в третьем классе должны оказаться все остальные 10 испытуемых. Мы ограничены одним условием - объемом выборки. Поэтому даже если мы потеряли данные о том, сколько человек не умеют работать на компьютере, мы можем определить это, зная, что в первом и втором классах - по 20 испытуемых. Мы не свободны в определении количества испытуемых в третьем разряде, "свобода" простирается только на первые две ячейки классификации:

$\nu = c-1 = 3-1 = 2$

Аналогичным образом, если бы у нас была классификация из 10 разрядов, то мы были бы свободны только в 9 из них, если бы у нас было 100 классов - то в 99 из них и т. д.

Способы более сложного подсчета числа степеней свободы при двухмерных классификациях приведены в разделах, посвященных критерию χ^2 и дисперсионному анализу.

Зная n и/или число степеней свободы, мы по специальным таблицам можем определить критические значения критерия и сопоставить с ними полученное эмпирическое значение. Обычно это записывается так: "при n=22 критические значения критерия составляют ..." или "при ν=2 критические значения критерия составляют ..." и т.п.

Критерии делятся на параметрические и непараметрические.

Параметрические критерии

Критерии, включающие в формулу расчета параметры распределения, то есть средние и дисперсии (*t* - критерий Стьюдента, критерий F и др.)

Непараметрические критерии

Критерии, не включающие в формулу расчета параметров распределения и основанные на оперировании частотами или рангами (критерий Q Розенбаума, критерий T Вилкоксона и др.)

И те, и другие критерии имеют свои преимущества и недостатки. На основании нескольких руководств можно составить таблицу, позволяющую оценить возможности и ограничения тех и других (Рунион Р., 1982; McCall R., 1970; J.Greene, M.D'Olivera, 1989).

Таблица 1.1

Возможности и ограничения параметрических и непараметрических критериев

ПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ	НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ
1. Позволяют прямо оценить различи* в средних, полученных в двух выборках (t - критерий Стьюдента).	Позволяют оценить лишь средние тенденции, например, ответить на вопрос, чаще ли в выборке А встречаются более высокие, а в выборке Б - более низкие значения признака (критерии Q, U, φ* и др.).
2. Позволяют прямо оценить различия в дисперсиях (критерий Фишера).	Позволяют оценить лишь различия в диапазонах вариативности признака (критерий φ*).
3. Позволяют выявить тенденции изменения признака при переходе от условия к условию (дисперсионный однофакторный анализ), но лишь при условии нормального распределения признака.	Позволяют выявить тенденции изменения признака при переходе от условия к условию при любом распределении признака (критерии тенденций L и S).
4. Позволяют оценить взаимодействие двух и более факторов в их влиянии на изменения признака (двухфакторный дисперсионный анализ).	Эта возможность отсутствует.

5. Экспериментальные данные должны отвечать двум, а иногда трем, условиям: а) значения признака измерены по интервальной шкале; б) распределение признака является нормальным; в) в дисперсионном анализе должно соблюдаться требование равенства дисперсий в ячейках комплекса.	Экспериментальные данные могут не отвечать ни одному из этих условий: а) значения признака могут быть представлены в любой шкале, начиная от шкалы наименований; б) распределение признака может быть любым и совпадение его с каким-либо теоретическим законом распределения необязательно и не нуждается в проверке; в) требование равенства дисперсий отсутствует.
6. Математические расчеты довольно сложны.	Математические расчеты по большей части просты и занимают мало времени (за исключением критериев χ^2 и λ).
7. Если условия, перечисленные в п.5, выполняются, параметрические критерии оказываются несколько более мощными, чем непараметрические.	Если условия, перечисленные в п.5, не выполняются, непараметрические критерии оказываются более мощными, чем параметрические, так как они менее чувствительны к "засорениям'.

Из Табл. 1.1 мы видим, что параметрические критерии могут оказаться несколько более мощными[5], чем непараметрические, но только в том случае, если признак измерен по интервальной шкале и нормально распределен. С интервальной шкалой есть определенные проблемы (см. раздел "Шкалы измерения"). Лишь с некоторой натяжкой мы можем считать данные, представленные не в стандартизованных оценках, как интервальные. Кроме того, проверка распределения "на нормальность" требует достаточно сложных расчетов, результат которых заранее неизвестен (см. параграф 7.2). Может оказаться, что распределение признака отличается от нормального, и нам так или иначе все равно придется обратиться к непараметрическим критериям.

Непараметрические критерии лишены всех этих ограничений и нетребуют таких длительных и сложных расчетов. По сравнению с параметрическими критериями они ограничены лишь в одном - с их помощью невозможно оценить взаимодействие двух или более условий или факторов, влияющих на изменение признака. Эту задачу может решить только дисперсионный двухфакторный анализ.

Учитывая это, в настоящее руководство включены в основном непараметрические статистические критерии. В сумме они охватывают большую часть возможных задач сопоставления данных.

Единственный параметрический метод, включенный в руководство - метод дисперсионного анализа, двухфакторный вариант которого ничем невозможно заменить.

1.6. Уровни статистической значимости

Уровень значимости - это вероятность того, что мы сочли различия существенными, а они на самом деле случайны.

Когда мы указываем, что различия достоверны на 5%-ом уровне значимости, или при $p<0,05$, то мы имеем виду, что вероятность того, что они все-таки недостоверны, составляет 0,05.

Когда мы указываем, что различия достоверны на 1%-ом уровне значимости, или при $p<0,01$, то мы имеем в виду, что вероятность того, что они все-таки недостоверны, составляет 0,01.

Если перевести все это на более формализованный язык, то уровень значимости - это вероятность отклонения нулевой гипотезы, в то время как она верна.

Ошибка, состоящая в том, что мы отклонили нулевую гипотезу, в то время как

[5] О понятии мощности критерия см. ниже.

она верна, называется ошибкой 1 рода.

Вероятность такой ошибки обычно обозначается какα. В сущности, мы должны были бы указывать в скобках не $p \leq 0{,}05$ или $p \leq 0{,}01$, а $\alpha \leq 0{,}05$ или $\alpha \leq 0{,}01$. В некоторых руководствах так и делается (Рунион Р., 1982; Захаров В.П., 1985 и др.).

Если вероятность ошибки - это α, то вероятность правильного решения: 1 — α. Чем меньше α, тем больше вероятность правильного решения.

Исторически сложилось так, что в психологии принято считать низшим уровнем статистической значимости 5%-ый уровень ($p<0{,}05$): достаточным - 1%-ый уровень ($p<0{,}01$) и высшим 0,1%-ый уровень ($p<0{,}001$), поэтому в таблицах критических значений обычно приводятся значения критериев, соответствующих уровням статистической значимости $p<0{,}05$ и $p<0{,}01$, иногда - $p<0{,}001$. Для некоторых критериев в таблицах указан точный уровень значимости их разных эмпирических значений. Например, для $\varphi^*=1{,}56$ $p=0{,}06$.

До тех пор, однако, пока уровень статистической значимости не достигнет $p=0{,}05$, мы еще не имеем права отклонить нулевую гипотезу. В настоящем руководстве мы, вслед за Р. Рунионом (1982), будем придерживаться следующего правила отклонения гипотезы об отсутствии различий (H_0) и принятия гипотезы о статистической достоверности различий (H_1).

Правило отклонения H_0 и принятия H_1

Если эмпирическое значение критерия равняется критическому значению, соответствующему $p<0{,}05$ или превышает его, то H_0 отклоняется, но мы еще не можем определенно принять H_1. Если эмпирическое значение критерия равняется критическому значению, соответствующему $p<0{,}01$ или превышает его, то H_0 отклоняется и принимается H_1.

Исключения: критерий знаков G, критерий T Вилкоксона и критерий U Манна-Уитни. Для них устанавливаются обратные соотношения.

Для облегчения процесса принятия решения можно всякий раз вычерчивать "ось значимости".



Рис. 1.7. Пример "оси значимости" для критерия Q Розенбаума

Критические значения критерия обозначены как $Q_{0,05}$ и $Q_{0,01}$, эмпирическое значение критерия как $Q_{эмп}$. Оно заключено в эллипс.

Вправо от критического значения $Q_{0,01}$ простирается "зона значимости" - сюда попадают эмпирические значения, превышающие $Q_{0,01}$ и, следовательно, безусловно значимые.

Влево от критического значения $Q_{0,05}$ простирается "зона незначимости", - сюда попадают эмпирические значения Q, которые ниже $Q_{0,05}$, и, следовательно, безусловно незначимы.

Мы видим, что $Q_{0,05}=6$; $Q_{0,01}=9$; $Q_{эмп}=8$

Эмпирическое значение критерия попадает в область между $Q_{0,05}$ и $Q_{0,01}$- Это зона "неопределенности": мы уже можем отклонить гипотезу о недостоверности различий (H_0), но еще не можем принять гипотезы об их достоверности (H_1).

Практически, однако, исследователь может считать достоверными уже те различия, которые не попадают в зону незначимости, заявив, что они достоверны при $p<0{,}05$, или

указав точный уровень значимости полученного эмпирического значения критерия, например: p=0,02. С помощью таблиц Приложения 1 это можно сделать по отношению к критериям H Крускала-Уоллиса, χ^2, Фридмана, L Пейджа, φ* Фишера, λ, Колмогорова.

Уровень статистической значимости или критические значения критериев определяются по-разному при проверке направленных и ненаправленных статистических гипотез.

При направленной статистической гипотезе используется односторонний критерий, при ненаправленной гипотезе - двусторонний критерий. Двусторонний критерий более строг, поскольку он проверяет различия в обе стороны, и поэтому то эмпирическое значение критерия, которое ранее соответствовало уровню значимости $p<0,05$, теперь соответствует лишь уровню $p<0,10$.

В данном руководстве исследователю не придется всякий раз самостоятельно решать, использует ли он односторонний или двухсторонний критерий. Таблицы критических значений критериев подобраны таким образом, что направленным гипотезам соответствует односторонний, а ненаправленным - двусторонний критерий, и приведенные значения удовлетворяют тем требованиям, которые предъявляются к каждому из них. Исследователю необходимо лишь следить за тем, чтобы его гипотезы совпадали по смыслу и по форме с гипотезами, предлагаемыми в описании каждого из критериев.

1.7. Мощность критериев

Мощность критерия - это его способность выявлять различия, если они есть. Иными словами, это его способность отклонить нулевую гипотезу об отсутствии различий, если она неверна.

Ошибка, состоящая в том, что мы приняли нулевую гипотезу, в то время как она неверна, называется ошибкой II рода.

Вероятность такой ошибки обозначается как β. Мощность критерия - это его способность не допустить ошибку II рода, поэтому:

Мощность=1—β

Мощность критерия определяется эмпирическим путем. Одни и те же задачи могут быть решены с помощью разных критериев, при этом обнаруживается, что некоторые критерии позволяют выявить различия там, где другие оказываются неспособными это сделать, или выявляют более высокий уровень значимости различий. Возникает вопрос: а зачем же тогда использовать менее мощные критерии? Дело в том, что основанием для выбора критерия может быть не только мощность, но и другие его характеристики, а именно:

а) простота;

б) более широкий диапазон использования (например, по отношению к данным, определенным по номинативной шкале, или по отношению к большим n);

в) применимость по отношению к неравным по объему выборкам;

г) большая информативность результатов.

1.8. Классификация задач и методов их решения

Множество задач психологического исследования предполагает те или иные сопоставления. Мы сопоставляем группы испытуемых по какому-либо признаку, чтобы выявить различия между ними по этому признаку. Мы сопоставляем то, что было "до" с тем, что стало "после" наших экспериментальных или любых иных воздействий, чтобы определить эффективность этих воздействий. Мы сопоставляем эмпирическое распределение значений признака с каким-либо теоретическим законом распределения или два эмпирических распределения между собой, с тем, чтобы доказать неслучайность

выбора альтернатив или различия в форме распределений.

Мы, далее, можем сопоставлять два признака, измеренные на одной и той же выборке испытуемых, для того, чтобы установить степень согласованности их изменений, их сопряженность, корреляцию между ними.

Наконец, мы можем сопоставлять индивидуальные значения, полученные при разных комбинациях каких-либо существенных условий, с тем чтобы выявить характер взаимодействия этих условий в их влиянии на индивидуальные значения признака.

Именно эти задачи позволяет решить тот набор методов, который предлагается настоящим руководством. Все эти методы могут быть использованы при так называемой "ручной" обработке данных.

Краткая классификация задач и методов дана в Таблице 1.2.

Таблица 1.2

Классификация задач и методов их решения

Задачи	**Условия**	**Методы**
1. Выявление различий в уровне исследуемого признака	а) 2 выборки испытуемых	**Q** - критерий Розенбаума; **U** - критерий Манна-Уитни; φ^* - критерий (угловое преобразование Фишера)
	б) 3 и более выборок испытуемых	**S** - критерий тенденций Джонкира; **H** - критерий Крускала-Уоллиса.
2. Оценка сдвига значений исследуемого признака	а) 2 замера на одной и той же выборке испытуемых	**T** - критерий Вилкоксона; **G** - критерий знаков; φ^* - критерий (угловое преобразование Фишера).
	б) 3 и более замеров на одной и той же выборке испытуемых	$\chi_л^2$ - критерий Фридмана; **L** - критерий тенденций Пейджа.
3. Выявление различий в распределении	а) при сопоставлении эмпирического признака распределения с теоретическим	χ^2 - критерий Пирсона; λ - критерий Колмогорова-Смирнова; **m** - биномиальный критерий.
	б) при сопоставлении двух эмпирических распределений	χ^2 - критерий Пирсона; λ - критерий Колмогорова-Смирнова; φ^* - критерий (угловое преобразование Фишера).
4.Выявление степени согласованности изменений	а) двух признаков	r_s - коэффициент ранговой корреляции Спирмена.
	б) двух иерархий или профилей	r_s - коэффициент ранговой корреляции Спирмена.
5. Анализ изменений признака под влиянием контролируемых условий	а) под влиянием одного фактора	**S** - критерий тенденций Джонкира; **L** - критерий тенденций Пейджа; однофакторный дисперсионный анализ Фишера.
	б) под влиянием двух факторов одновременно	Двухфакторный дисперсионный анализ Фишера.

1.9. Принятие решения о выборе метода математической обработки

Если данные уже получены, то вам предлагается следующий алгоритм определения задачи и метода.

АЛГОРИТМ 1

Принятие решения о задаче и методе обработки на стадии, когда данные уже получены

1. По первому столбцу Табл. 1.2 определить, какая из задач стоит в вашем

исследовании.

2. По второму столбцу Табл. 1.2 определить, каковы условия решения вашей задачи, например, сколько выборок обследовано или на какое количество групп вы можете разделить обследованную выборку.

3. Обратиться к соответствующей главе и по алгоритму принятия решения о выборе критерия, приведенного в конце каждой главы, определить, какой именно метод или критерий вам целесообразно использовать.

Если вы еще находитесь на стадии планирования исследования, то лучшее заранее подобрать математическую модель, которую вы будете в дальнейшем использовать. Особенно необходимо планирование в тех случаях, когда в перспективе предполагается использование критериев тенденций или (в еще большей степени) дисперсионного анализа. , В этом случае алгоритм принятия решения таков:

АЛГОРИТМ 2

Принятие решения о задаче и методе обработки на стадии планирования исследования

1. Определите, какая модель вам кажется наиболее подходящей для доказательства] ваших научных предположений.

2. Внимательно ознакомьтесь с описанием метода, примерами и задачами для самостоятельного решения, которые к нему прилагаются.

3. Если вы убедились, что это то, что вам нужно, вернитесь к разделу "Ограничения критерия" и решите, сможете ли вы собрать данные, которые будут отвечать этим ограничениям (большие объемы выборок, наличие нескольких выборок, монотонно различающихся по какому-либо признаку, например, по возрасту и т.п.).

4. Проводите исследование, а затем обрабатывайте полученные данные по заранее! выбранному алгоритму, если вам удалось выполнить ограничения.

5. Если ограничения выполнить не удалось, обратитесь к алгоритму 1.

В описании каждого критерия сохраняется следующая последовательность изложения:

- назначение критерия;
- описание критерия;
- гипотезы, которые он позволяет проверить;
- графическое представление критерия;
- ограничения критерия;
- пример или примеры.

Кроме того, для каждого критерия создан алгоритм расчетов. Если критерий сразу удобнее рассчитывать по алгоритму, то он приводится в разделе "Пример"; если алгоритм легче можно воспринять уже после рассмотрения примера, то он приводится в конце параграфа, соответствующего данному критерию.

1.10. Список обозначений

Латинские обозначения:

A - показатель асимметрии распределения
c - количество групп или условий измерения
d - разность между рангами, частотами или частостями
df - число степеней свободы в дисперсионном анализе
E - показатель эксцесса
F - критерий Фишера для сравнения дисперсий
f - частота
*f** - частость, или относительная частота

G - критерий знаков
H - критерий Крускала-Уоллиса
i - индекс, обозначающий порядковый номер наблюдения
j - индекс, обозначающий порядковый номер разряда, класса, группы
k - количество классов или разрядов признака
L - критерий тенденций Пейджа
M - среднее значение признака или средняя арифметическая; то же, что и $\bar{x}$
m - биномиальный критерий
n - количество наблюдений (испытуемых, реакций, выборов и т.п.)
N - общее количество наблюдений в двух или более выборках
P - вероятность того, что событие произойдет
p - вероятность ошибки 1 рода (то же, что и а), уровень статистической значимости
Q - 1) вероятность того, что событие не произойдет; 2) критерий Розенбаума
r_s - коэффициент ранговой корреляции Спирмена
S - критерий Джонкира
S^2 - оценка дисперсии
S_i - количество значений, которые выше или ниже данного значения
SS - суммы квадратов (в дисперсионном анализе)
T - критерий Вилкоксона
T_c - суммы рангов по столбцам
$T_к$ - большая сумма рангов в критерии U
U - критерий Манна-Уитни
Wn - размах вариативности, или диапазон значений от наименьшего до наибольшего
x_i - текущее наблюдение; каждое наблюдение по порядку
$\bar{x}$ - среднее значение признака (то же, что и M)

Греческие обозначения:

α (альфа)	- вероятность ошибки I рода (отклонения H_0, которая верна)
β (бета)	- вероятность ошибки II рода (принятия H_0, которая неверна)
λ, (ламбда)	- критерий Колмогорова-Смирнова
ν (ню)	- число степеней свободы в непараметрических критериях
σ (сигма)	- стандартное отклонение
φ (фи)	- центральный угол, определяемый по процентной доле в критерии φ^*
φ^* (фи)	- критерий Фишера с угловым преобразованием
χ^2 (хи-квадрат)	- критерий Пирсона
χ^2_r (хи-ар-квадрат)	- критерий Фридмана.

ГЛАВА 2
ВЫЯВЛЕНИЕ РАЗЛИЧИЙ В УРОВНЕ ИССЛЕДУЕМОГО ПРИЗНАКА

2.1. Обоснование задачи сопоставления и сравнения

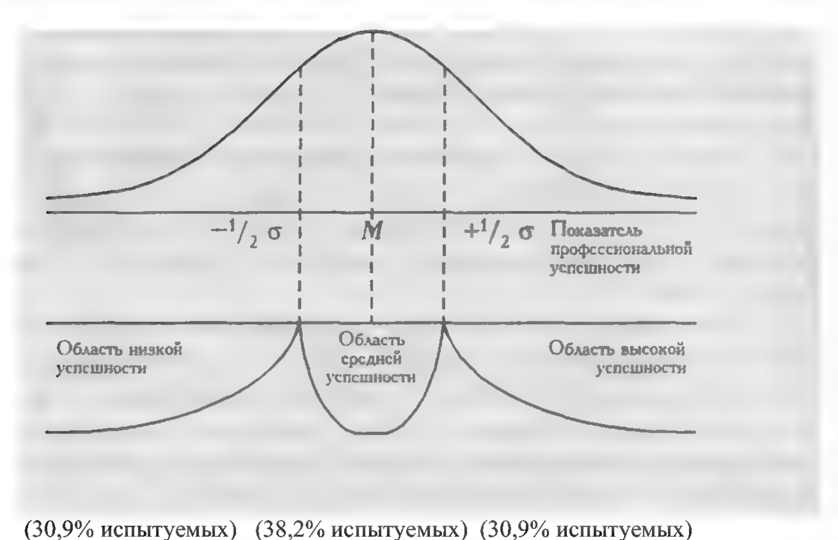
Очень часто перед исследователем в психологии стоит задача выявления различий между двумя, тремя и более выборками испытуемых. Это может быть, например, задача определения психологических особенностей хронически больных детей по сравнению со здоровыми, юных правонарушителей по сравнению с законопослушными сверстниками или различий между работниками государственных предприятий и частных фирм, между людьми разной национальности или разной культуры и, наконец, между людьми разного возраста в методе "поперечных срезов".

Иногда по выявленным в исследовании статистически достоверным различиям формируется "групповой профиль" или "усредненный портрет" человека той или иной профессии, статуса, соматического заболевания и др. (см., например, Cattell R.B., Eber H.W., Tatsuoka M.M., 1970).

В последние годы все чаще встает задача выявления психологического портрета специалиста новых профессий: "успешного менеджера", "успешного политика", "успешного торгового представителя", "успешного коммерческого директора" и др. Такого рода исследования не всегда подразумевают участие двух или более выборок. Иногда обследуется одна, но достаточно представительная выборка численностью не менее 60 человек, а затем внутри, этой выборки выделяются группы более и менее успешных специалистов, и их данные по исследованным переменным сопоставляются между собой. В самом простом случае критерием для разделения выборки на "успешных" и "неуспешных" будет средняя величина по показателю успешности. Однако такое деление является довольно грубым: лица, получившие близкие оценки по успешности, могут оказаться в противоположных группах, а лица, заметно различающиеся по оценкам успешности, - в одной и той же группе.

Это может исказить результаты сопоставления групп или по крайней мере сделать различия между группами менее заметными.

Чтобы избежать этого, можно попробовать выделить группы "успешных" и "неуспешных" специалистов более строго, включая в первую из них только тех, чьи значения *превышают* среднюю величину не менее чем на 1/4 стандартного отклонения, а во вторую группу - только тех, чьи значения не менее чем на 1/4 стандартного отклонения *ниже* средней величины. При этом все, кто оказывается в зоне средних величин, $M \pm 1/4\ \sigma$, выпадают из дальнейших сопоставлений. Если распределение близко к нормальному, то выпадет примерно 19,8% испытуемых. Если распределение отличается от нормального, то таких испытуемых может быть и больше. Чтобы избежать потерь, можно сопоставлять не две, а три группы испытуемых: с высокой, средней и низкой профессиональной успешностью.



(30,9% испытуемых) (38,2% испытуемых) (30,9% испытуемых)

Рис 2.1. Схематическое изображение процесса разделения выборки на группы с низкой, средней и высокой профессиональной успешностью

На Рис. 2.1 представлена схема разделения выборки на группы с низкой, средней и высокой профессиональной успешностью по критерию отклонения значений от средней величины на 1/2 стандартного отклонения. При таком строгом критерии в "среднюю" группу попадают (при нормальном распределении) около 38,2% всех испытуемых, а в крайних группах оказывается по 30,9% испытуемых.

Чем меньше испытуемых оказывается в группах, тем меньше у нас возможностей для выявления достоверных различий, так как критические значения большинства критериев при малых *n* строже, чем при больших *n*.

Таким образом, при нестрогом разделении испытуемых на группы мы теряем в точности, а при строгом - в количестве испытуемых.

При решении задач выявления различий в уровневых показателях следует помнить, что "усредненный профиль успешного специалиста" должен рассматриваться скорее как исследовательский результат, позволяющий сформулировать гипотезы для дальнейших исследований, а не как основание для профессионального отбора. Тому есть две причины. Во-первых, ни у одного из успешных специалистов может не наблюдаться "усредненный профиль" - он, в сущности, является отвлеченным обобщением; во-вторых, в профессиональной деятельности наличие собственного индивидуального стиля важнее соответствия "среднегрупповому" профилю. Недостаток в тех качествах, которые могут казаться важными, компенсируется другими качествами. У каждого успешного специалиста его психологические свойства создают неповторимый ансамбль, который при усреднении данных теряется.

Р.Б. Кеттелл, учитывая это, предлагал при исследовании профессиональной успешности включать в рассмотрение индивидуальные профили выдающихся представителей той или иной профессии (Cattell R.B., Eber H.W., Tatsuoka M.M., 1970).

Сопоставление уровневых показателей в разных выборках может быть необходимой частью комплексных диагностических, учебных, пси-хокоррекционных и иных программ. Оно помогает нам обратить внимание на те особенности обследованных выборок, которые должны быть учтены и использованы при адаптации программ к данной группе в процессе их конкретного воплощения.

Критерии, которые рассматриваются в данной главе, предполагают, что мы сопоставляем так называемые независимые выборки, то есть две или более выборки, состоящие из разных испытуемых. Тот испытуемый, который входит в одну выборку, уже не может входить в другую. В противоположность этому, если мы обследуем одну и ту же выборку испытуемых, несколько раз подвергая ее аналогичным измерениям ("замерам"), то перед нами - так называемые связанные, или зависимые, выборки данных.

Сопоставление 2-х или более замеров, полученных на одной и той же выборке, рассматривается в Главе 3.

Решение о выборе того или иного критерия принимается на основе того, сколько выборок сопоставляется и каков их объем (см. Алгоритм 7 в конце главы).

2.2. Q - критерий Розенбаума

Назначение критерия

Критерий используется для оценки различий между *двумя* выборками *по уровню* какого-либо признака, количественно измеренного. В каждой из выборок должно быть не менее 11 испытуемых.

Описание критерия

Это очень простой непараметрический критерий, который позволяет быстро оценить различия между двумя выборками по какому-либо признаку. Однако если критерий Q не выявляет достоверных различий, это еще не означает, что их действительно нет.

В этом случае стоит применить критерий φ^* Фишера. Если же Q-критерий выявляет достоверные различия между выборками с уровнем значимости $p<0,01$, можно ограничиться только им и избежать трудностей применения других критериев.

Критерий применяется в тех случаях, когда данные представлены по крайней мере в порядковой шкале. Признак должен варьировать в каком-то диапазоне значений, иначе сопоставления с помощью Q -критерия просто невозможны. Например, если у нас только 3 значения признака, 1, 2 и 3, - нам очень трудно будет установить различия. Метод Розенбаума требует, следовательно, достаточно тонко измеренных признаков.

Применение критерия начинаем с того, что упорядочиваем значения признака в обеих выборках по нарастанию (или убыванию) признака. Лучше всего, если данные каждого испытуемого представлены на отдельной карточке. Тогда ничего не стоит упорядочить два ряда значений по интересующему нас признаку, раскладывая карточки на столе. Так мы сразу увидим, совпадают ли диапазоны значений, и если нет, то насколько один ряд значений "выше" (S_1), а второй - "ниже" (S_2). Для того, чтобы не запутаться, в этом и во многих других критериях рекомендуется первым рядом (выборкой, группой) считать тот ряд, где значения выше, а вторым рядом - тот, где значения ниже.

Гипотезы

H_0: Уровень признака в выборке 1 не превышает уровня признака в выборке 2.

H_1: Уровень признака в выборке 1 превышает уровень признака в выборке 2.

Графическое представление критерия Q

На Рис. 2.2. представлены три варианта соотношения рядов значений в двух выборках. В варианте (а) все значения первого ряда выше всех значений второго ряда. Различия, безусловно, достоверны, при соблюдении условия, что n_1, $n_2\geq 11$.

В варианте (б), напротив, оба ряда находятся на одном и том же уровне: различия недостоверны. В варианте (в) ряды частично перекрещиваются, но все же первый ряд оказывается гораздо выше второго. Достаточно ли велики зоны S_1 и S_2, в сумме составляющие Q, можно определить по Таблице I Приложения 1, где приведены критические значения Q для разных *n*. Чем величина Q больше, тем более достоверные различия мы сможем констатировать.

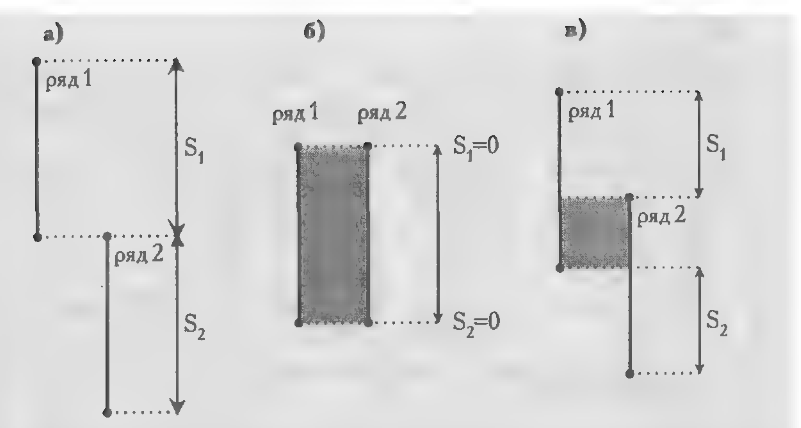


Рис. 2.2. Возможные соотношения рядов значений в двух выборках; S_1 - зона значений 1-го ряда, которые выше максимального значения 2-го ряда; S_2 - зона значений второго ряда, которые меньше минимального значения 1-го ряда; штриховкой отмечены перекрещивающиеся зоны двух рядов

Ограничения критерия Q

1. В каждой из сопоставляемых выборок должно быть не менее 11 наблюдений. При этом объемы выборок должны примерно совпадать. Е.В. Гублером указываются следующие правила:

а) если в обеих выборках меньше 50 наблюдений, то абсолютная величина разности между n_1 и n_2 не должна быть больше 10 наблюдений;

б) если в каждой из выборок больше 51 наблюдения, но меньше 100, то абсолютная величина разности между *щ* и Л2 не должна быть больше 20 наблюдений;

в) если в каждой из выборок больше 100 наблюдений, то допускается, чтобы одна из выборок была больше другой не более чем в 1,5-2 раза (Гублер Е.В., 1978, с. 75).

2. Диапазоны разброса значений в двух выборках должны не совпадать между собой, в противном случае применение критерия бессмысленно. Между тем, возможны случаи, когда диапазоны разброса значений совпадают, но, вследствие разносторонней асимметрии двух распределений, различия в средних величинах признаков существенны (Рис. 2.3., 2.4).

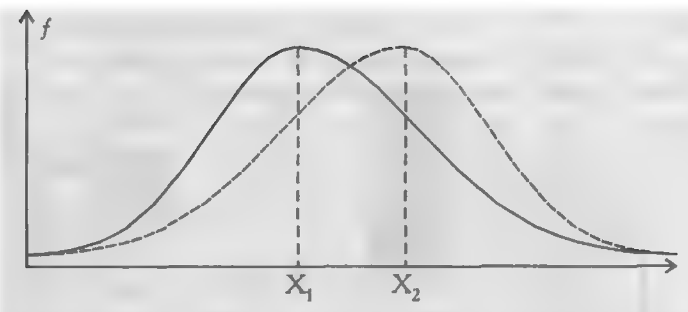


Рис. 2.3. Вариант соотношения распределений признака в двух выборках, при котором критерий Q беспомощен

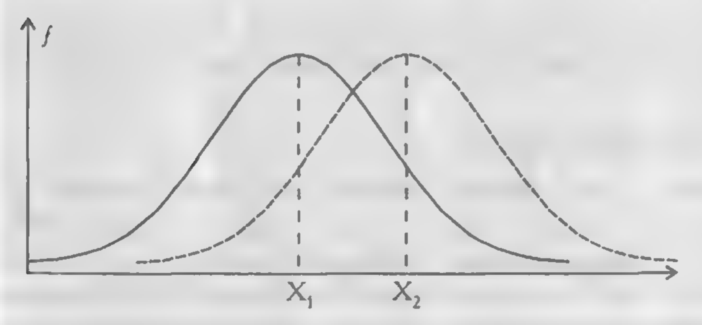


Рис. 2.4. Вариант соотношения распределений признака в двух выборках, при котором критерий Q может быть могущественным

Пример

У предполагаемых участников психологического эксперимента, моделирующего

деятельность воздушного диспетчера, был измерен уровень вербального и невербального интеллекта с помощью методики Д. Векслера. Было обследовано 26 юношей в возрасте от 18 до 24 лет (средний возраст 20,5 лет). 14 из них были студентами физического факультета, а 12 - студентами психологического факультета Ленинградского университета (Сидоренко Е.В., 1978). Показатели вербального интеллекта представлены в Табл. 2.1.

Можно ли утверждать, что одна из групп превосходит другую по уровню вербального интеллекта?

Таблица 2.1

Индивидуальные значения вербального интеллекта в выборках студентов физического (n_1=14) и психологического (n_2=12) факультетов

Студе нты-физики			Студенты - психологи		
	Код имени испытуемого	Показатель вербального интеллекта		Код имени испытуемого	Показатель вербального интеллекта
1.	И.А	132	1.	н.т.	126
2.	К.А.	134	2.	о.в.	127
3.	К.Е.	124	3.	Е.В.	132
4.	П.А.	132	4.	Ф.О.	120
5.	С.А.	135	5.	и.н.	119
6.	Ст.А.	132	6.	и.ч.	126
7.	Т.А.	131	7.	и.в.	120
8.	Ф.А.	132	8.	К.О.	123
9.	Ч.И.	121	9.	р.р.	120
10.	Ц.А.	127	10	Р.И.	116
11.	См.А.	136	11	О.К.	123
12.	К.Ан.	129	12	н.к.	115
13.	Б.Л.	136			
14.	Ф.В.	136			

Упорядочим значения в обеих выборках, а затем сформулируем гипотезы:

H_0: Студенты-физики не превосходят студентов-психологов по уровню вербального интеллекта.

H_1: Студенты-физики превосходят студентов-психологов по уровню вербального интеллекта.

Таблица 2.2

Упорядоченные по убыванию вербального интеллекта ряды индивидуальных значений в двух студенческих выборках

1 ряд - студенты-физики					2 ряд - студенты-психологи		
1	См.А.	136					
2	Б.Л.	136					
3	Ф.В.	136	S_1				
4	С.А.	135					
5	К.А.	134					
6	И.А.	132			1	Е.В.	132
7	П.А.	132					
8	Ст.А.	132					
9	Ф.А.	132					
10	Т.А.	131					
11	К.Ан.	129					
12	Ц.А.	127			2	О.В.	127
					3	Н.Т.	126
					4	И.Ч.	126
13	К.Е.	124					
					5	К.О.	123
					6	О.К.	123
14	Ч.И.	121					
					7	Ф.О.	120
					8	И.В.	120
				S_2	9	Р.Р.	120
					10	И.Н.	119
					11	Р.И.	116
					12	Н.К.	115

Как видно из Табл. 2.2, мы правильно обозначили ряды: первый, тот, что "выше" - ряд физиков, а второй, тот, что "ниже" - ряд психологов.

По Табл. 2.2 определяем количество значений первого ряда, которые больше максимального значения второго ряда: $S_1=5$.

Теперь определяем количество значений второго ряда, которые меньше минимального значения первого ряда: $S_2=6$.

Вычисляем $Q_{ЭМП}$ по формуле:

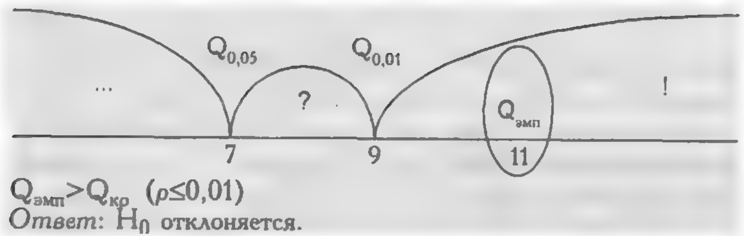
$Q_{ЭМN}=S_1+S_2=5+6=ll$

По Табл.II Приложения 1 определяем критические значения Q для $n_1=14$, $n_2=12$:

$$Q_{кр}=\begin{cases} 7 & (p \le 0{,}05) \\ 9 & (p \le 0{,}01) \end{cases}$$

Ясно, что чем больше расхождения между выборками, тем больше величина Q. H_o отклоняется при $Q_{эмп}>Q_{кр}$, а при $Q_{эмп}<Q_{кр}$ мы будем вынуждены принять H_o.

Построим "ось значимости".



Принимается H_1. Студенты-физики превосходят студентов-психологов по уровню вербального интеллекта ($p<0{,}01$). Отметим, что в тех случаях, когда эмпирическая величина критерия оказывается на границе зоны незначимости, мы имеем право утверждать лишь, что различия достоверны при $p<0{,}05$, если же оно оказывается между двумя критическими значениями, то мы можем утверждать, что $p<0{,}05$.

Если эмпирическое значение критерия оказывается на границе зоны значимости, $p<0{,}01$, в зоне значимости - что $p<0{,}01$

Поскольку уровень значимости выявленных различий достаточно высок ($p<0{,}01$), мы могли бы на этом остановиться. Однако если исследователь сам психолог, а не физик, вряд ли он на этом остановится. Он может попробовать сопоставить выборки по уровню невербального интеллекта, поскольку именно невербальный интеллект определяет уровень интеллекта в целом и степень его организованности (см., например: Бергер М.А., Логинова НА., 1974).

Мы вернемся к этому примеру при рассмотрении критерия Манна-Уитни и

попытаемся ответить на вопрос о соотношении уровней невербального интеллекта в двух выборках. Быть может, психологи еще окажутся в более высоком ряду!

АЛГОРИТМ 3
Подсчет критерия Q Розенбаума

1. Проверить, выполняются ли ограничения: $n_1 \bullet n_2 \geq 11$, n_1 $n_2 \approx n_2$Упорядочить значения отдельно в каждой выборке по степени возрастания признака. Считать выборкой 1 ту выборку, значения в которой предположительно выше, а выборкой 2 - ту, где значения предположительно ниже.

3. Определить самое высокое (максимальное) значение в выборке 2.

4. Подсчитать количество значений в выборке 1, которые выше максимального значения в выборке 2. Обозначить полученную величину как S_1.

5. Определить самое низкое (минимальное) значение в выборке 1.

6. Подсчитать количество значений в выборке 2, которые ниже минимального значения выборки 1. Обозначить полученную величину как S_2.

7. Подсчитать эмпирическое значение Q по формуле: $Q=S_1$+S2-

8. По Табл. I Приложения I определить критические значения Q для данных n_1 и n_2. Если $Q_{эмп}$ равно $Q_{0,05}$ или превышает его, H_0 отвергается.

9. При $n_1 \bullet n_2 > 26$ сопоставить полученное эмпирическое значение с $Q_{кр=}8$ ($p \leq 0,05$) и $Q_{кр=}10$($p \leq 0,01$). Если $Q_{эмп}$ превышает или по крайней мере равняется $Q_{кр=}8$, H_0 отвергается.

2.3. U - критерий Манна-Уитни

Назначение критерия

Критерий предназначен для оценки различий между *двумя* выборками по *уровню* какого-либо признака, количественно измеренного. Он позволяет выявлять различия между *малыми* выборками, когда $n_1 \bullet n_2 \geq 3$ или $n_1=2$, $n_2 \geq 5$, и является более мощным, чем критерий Розенбаума.

Описание критерия

Существует несколько способов использования критерия и несколько вариантов таблиц критических значений, соответствующих этим способам (Гублер Е. В., 1978; Рунион Р., 1982; Захаров В. П., 1985; McCall R., 1970; Krauth J., 1988).

Этот метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами. Мы помним, что 1-м рядом (выборкой, группой) мы называем тот ряд значений, в котором значения, по предварительной оценке, выше, а 2-м рядом - тот, где они предположительно ниже.

Чем меньше область перекрещивающихся значений, тем более вероятно, что различия достоверны. Иногда эти различия называют различиями в *расположении* двух выборок (Welkowitz J. et al., 1982).

Эмпирическое значение критерия U отражает то, насколько велика зона совпадения между рядами. Поэтому *чем меньше* $U_{эмп}$, *тем более* вероятно, что различия *достоверны.*

Гипотезы

H_0: Уровень признака в группе 2 не ниже уровня признака в группе 1.

H_1: Уровень признака в группе 2 ниже уровня признака в группе 1.

Графическое представление критерия U

На Рис. 2.5. представлены три из множества возможных вариантов соотношения двух рядов значений.

В варианте (а) второй ряд ниже первого, и ряды почти не перекрещиваются. Область наложения слишком мала, чтобы скрадывать различия между рядами. Есть шанс, что

различия между ними достоверны. Точно определить это мы сможем с помощью критерия U.

В варианте (б) второй ряд тоже ниже первого, но и область перекрещивающихся значений у двух рядов достаточно обширна. Она может еще не достигать критической величины, когда различия придется признать несущественными. Но так ли это, можно определить только путем точного подсчета критерия U.

В варианте (в) второй ряд ниже первого, но область наложения настолько обширна, что различия между рядами скрадываются.

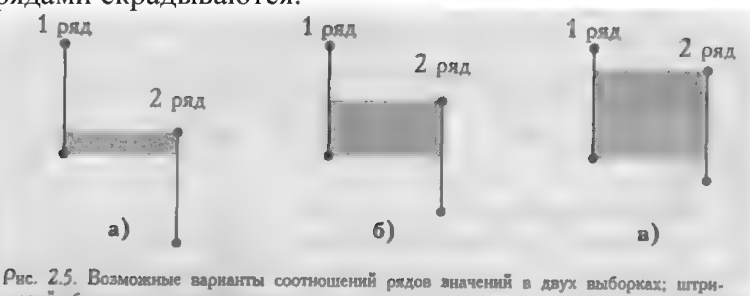


Рис. 2.5. Возможные варианты соотношений рядов значений в двух выборках; штриховкой обозначены зоны наложения

Ограничения критерия U

1. В каждой выборке должно быть не менее 3 наблюдений: $n_1 \bullet n_2 \geq 3$; допускается, чтобы в одной выборке было 2 наблюдения, но тогда во второй их должно быть не менее 5.

2. В каждой выборке должно быть не более 60 наблюдений; $n_1 \bullet n_2 \leq 60$. Однако уже при $n_1 \bullet n_2 > 20$ ранжирование становиться достаточно трудоемким.

На наш взгляд, в случае, если $n_1 \bullet n_2 > 20$, лучше использовать другой критерий, а именно угловое преобразование Фишера в комбинации с критерием λ,, позволяющим выявить критическую точку, в которой накапливаются максимальные различия между двумя сопоставляемыми выборками (см. п. 5.4). .Формулировка звучит сложно, но сам метод достаточно прост. Каждому исследователю лучше попробовать разные пути и выбрать тот, который кажется ему более подходящим.

Пример

Вернемся к результатам обследования студентов физического и психологического факультетов Ленинградского университета с помощью методики Д. Векслера для измерения вербального и невербального интеллекта. С помощью критерия Q Розенбаума мы в предыдущем параграфе смогли с высоким уровнем значимости определить, что уровень вербального интеллекта в выборке студентов физического факультета выше. Попытаемся установить теперь, воспроизводится ли этот результат при сопоставлении выборок по уровню невербального интеллекта. Данные приведены в Табл. 2.3.

Можно ли утверждать, что одна из выборок превосходит другую по уровню невербального интеллекта?

Таблица 2.3

Индивидуальные значения невербального интеллекта в выборках студентов физического *(щ=\4)* и психологического (п2=12) факультетов

Студенты-физики		Студенты-психологи	
Код имени испытуемого	Показатель невербального интеллекта	Код имени испытуемого	Показатель невербального интеллекта
1. И.А.	111	1. Н.Т.	ИЗ
2. К.А.	104	2. О.В.	107
3. К.Е.	107	3. Е.В.	123
4. П.А.	90	4. Ф.О.	122
5. С.А.	115	5. И.Н.	117
6. Ст.А.	107	6. И.Ч.	112
7. Т.А.	106	7. И.В.	105

8.	Ф.А.	107	8.	К.О.	108
9.	Ч.И.	95	9.	Р.Р.	111
10.	ЦА.	116	10.	Р.И.	114
11.	См.А.	127	11.	О.К.	102
12.	К.Ан.	115	12.	Н.К.	104
13.	Б.Л.	102			
14.	Ф.В.	99			

Критерий U требует тщательности и внимания. Прежде всего, необходимо помнить правила ранжирования.

Правила ранжирования

1. Меньшему значению начисляется меньший ранг. Наименьшему значению начисляется ранг 1.

Наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений. Например, если n=7, то наибольшее значение получит ранг 7, за возможным исключением для тех случаев, которые предусмотрены правилом 2.

2. В случае, если несколько значений равны, им начисляется ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны.

Например, 3 наименьших значения равны 10 секундам. Если бы мы измеряли время более точно, то эти значения могли бы различаться и составляли бы, скажем, 10,2 сек; 10,5 сек; 10,7 сек. В этом случае они получили бы ранги, соответственно, 1, 2 и 3. Но поскольку полученные нами значения равны, каждое из них получает средний ранг:

$$\frac{1+2+3}{3}=\frac{6}{3}=2$$

Допустим, следующие 2 значения равны 12 сек. Они должны были бы получить ранги 4 и 5, но, поскольку они равны, то получают средний ранг:

$$\frac{4+5}{2}=4.5 \text{ и т. д.}$$

3. Общая сумма рангов должна совпадать с расчетной, которая определяется по формуле:

$$\sum(R_i)=\frac{N\cdot(N+1)}{2}$$

где N - общее количество ранжируемых наблюдений (значений). Несовпадение реальной и расчетной сумм рангов будет свидетельствовать об ошибке, допущенной при начислении рангов или их суммировании. Прежде чем продолжить работу, необходимо найти ошибку и устранить ее.

При подсчете критерия U легче всего сразу приучить себя действовать по строгому алгоритму.

АЛГОРИТМ 4

Подсчет критерия U Манна-Уитни.

1. Перенести все данные испытуемых на индивидуальные карточки.

2. Пометить карточки испытуемых выборки 1 одним цветом, скажем красным, а все карточки из выборки 2 - другим, например синим.

3. Разложить все карточки в единый ряд по степени нарастания признака, не считаясь с тем, к какой выборке они относятся, как если бы мы работали с одной большой выборкой.

4. Проранжировать значения на карточках, приписывая меньшему значению меньший ранг. Всего рангов получится столько, сколько у нас (n_1+n_2).

5. Вновь разложить карточки на две группы, ориентируясь на цветные обозначения: красные карточки в один ряд, синие - в другой.

6. Подсчитать сумму рангов отдельно на красных карточках (выборка 1) и на синих карточках (выборка 2). Проверить, совпадает ли общая сумма рангов с расчетной.

7. Определить большую из двух ранговых сумм.

8. Определить значение U по формуле:

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x$$

где n_1 - количество испытуемых в выборке 1;

n_2 - количество испытуемых в выборке 2;

T_x - большая из двух ранговых сумм;

n_x - количество испытуемых в группе с большей суммой рангов.

9. Определить критические значения U по Табл. II Приложения 1. Если $U_{эмп.} > U_{кр\ 005}$, H_o принимается. Если $U_{эмп} \leq U_{кр\ 005}$, H_o отвергается. Чем меньше значения U, тем достоверность различий выше.

Теперь проделаем всю эту работу на материале данного примера. В результате работы по 1-6 шагам алгоритма построим таблицу.

Таблица 2.4

Подсчет ранговых сумм по выборкам студентов физического и психа-логического факультетов

Студенты-физики (n_1=14)			Студенты-психологи (n_2=12)	
Показатель невербального интеллекта		Ранг	Показатель невербального интеллекта	Ранг
	127	26		
			123	25
			122	24
			117	23
	116	22		
	115	20,5		
	115	20,5		
			114	19
			113	18
			112	17
	111	15,5	111	15.5
			108	14'
	107	11.5	107	11,5
	107	11,5		
	107	11,5		
	106	9		
			105	8
	104	6.5	104	6,5
	102	4,5	102	4,5
	99	3		
	95	2		
	90	1		
Суммы	1501	165	1338	186
Средние	107,2		111,5	

Общая сумма рангов: 165+186=351. Расчетная сумма:

$$\sum R_i = \frac{N \cdot (N+1)}{2} = \frac{26 \cdot (26+1)}{2} = 351$$

Равенство реальной и расчетной сумм соблюдено.

Мы видим, что по уровню невербального интеллекта более "высоким" рядом оказывается выборка студентов-психологов. Именно на эту выборку приходится большая ранговая сумма: 186.

Теперь мы готовы сформулировать гипотезы:

H_0: Группа студентов-психологов не превосходит группу студентов-физиков по уровню невербального интеллекта.

H_1: Группа студентов-психологов превосходит группу студентов-физиков по уровню невербального интеллекта.

В соответствии со следующим шагом алгоритма определяем эмпирическую

величину U:

$$U_{эмп} = (14 \cdot 12) + \frac{12 \cdot (12+1)}{2} - 186 = 60$$

Поскольку в нашем случае *n\Фn2,* подсчитаем эмпирическую величину U и для второй ранговой суммы (165), подставляя в формулу соответствующее ей n_x:

$$U_{эмп} = (14 \cdot 12) + \frac{14 \cdot (14+1)}{2} - 165 = 108$$

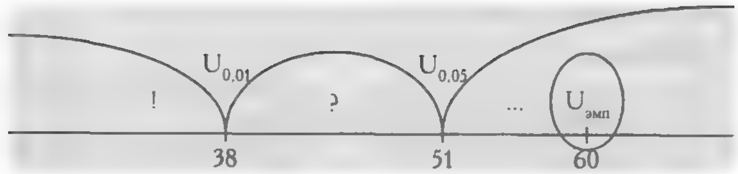
Такую проверку рекомендуется производить в некоторых руководствах (Рунион Р., 1982; Greene J., D'Olivera M., 1989). Для сопоставления с критическим значением выбираем меньшую величину U: $U_{эмп}=60$.

По Табл. II Приложения 1 определяем критические значения для $n_1=14$, $n_2=12$.

$$U_{кр} = \begin{cases} 51 & (p \le 0{,}05) \\ 38 & (p \le 0{,}01) \end{cases}$$

Мы помним, что критерий U является одним из двух исключений из общего правила принятия решения о достоверности различий, а именно, мы можем констатировать достоверные различия, если $U_{эмп} \le U_{кр}$

Построим "ось значимости".



$U_{эмп}=60$

$U_{эмп}>U_{кр}$

Ответ: H_0 принимается. Группа студентов-психологов не превосходит группы студентов-физиков по уровню невербального интеллекта.

Обратим внимание на то, что для данного случая критерий Q Розенбаума неприменим, так как размах вариативности в группе физиков шире, чем в группе психологов: и самое высокое, и самое низкое значение невербального интеллекта приходится на группу физиков (см. Табл. 2.4).

2.4. H - критерий Крускала-Уоллиса

Назначение критерия

Критерий предназначен для оценки различий одновременно между *тремя, четырьмя* и т.д. выборками по *уровню* какого-либо признака.

Он позволяет установить, что *уровень* признака *изменяется* при переходе от группы к группе, но не указывает на направление этих изменений.

Описание критерия

Критерий H иногда рассматривается как непараметрический аналог метода дисперсионного однофакторного анализа для несвязных выборок (Тюрин Ю. Н., 1978). Иногда его называют критерием "суммы рангов" (Носенко И.А., 1981).

Данный критерий является продолжением критерия U на большее, чем 2, количество сопоставляемых выборок. Все индивидуальные значения ранжируются так, как если бы это была одна большая выборка. Затем все индивидуальные значения возвращаются в свои первоначальные выборки, и мы подсчитываем суммы полученных ими рангов отдельно по каждой выборке. Если различия между выборками случайны, суммы рангов не будут различаться сколько-нибудь существенно, так как высокие и низкие ранги равномерно распределятся между выборками. Но если в одной из выборок будут преобладать низкие значения рангов, в другой - высокие, а в третьей - средние, то критерий H позволит установить эти различия.

Гипотезы

H_0: Между выборками 1, 2, 3 и т. д. существуют лишь случайные различия по уровню исследуемого признака.

H_1: Между выборками 1, 2, 3 и т. д. существуют неслучайные различия по уровню исследуемого признака.

Графическое представление критерия Н

Критерий Н оценивает общую сумму перекрещивающихся зон при сопоставлении всех обследованных выборок. Если суммарная область наложения мала (Рис. 2.6 (а)), то различия достоверны; если она достигает определенной критической величины и превосходит ее (Рис. 2.6 (б)), то различия между выборками оказываются недостоверными.

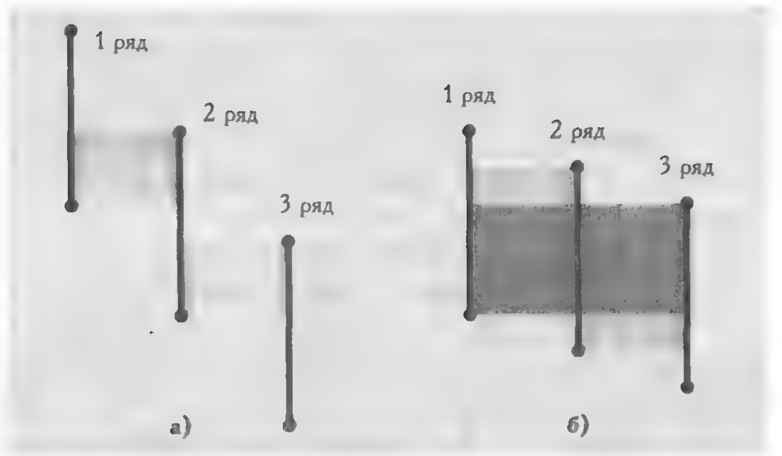


Рис. 2.6. 2 возможных варианта соотношения рядов значений в трех выборках; штриховкой отмечены зоны наложения

Ограничения критерия Н

1. При сопоставлении 3-х выборок допускается, чтобы в одной из них *n—Ъ,* а двух других n=2. Но при таких численных составах выборок мы сможем установить различия лишь на низшем уровне значимости ($p \leq 0,05$).

   Для того, чтобы оказалось возможным диагностировать различия на более высоком уровнем значимости (р5~0,01), необходимо, чтобы в каждой выборке было не менее 3 наблюдений, или чтобы по крайней мере в одной из них было 4 наблюдения, а в двух других - по 2; при этом неважно, в какой именно выборке сколько испытуемых, а важно соотношение 4:2:2.
2. Критические значения критерия Н и соответствующие им уровни значимости приведены в Табл. IV Приложения 1. Таблица предусмотрена только для трех выборок и $\{n_1, n_2, n_3\} \leq 5$.

   При большем количестве выборок и испытуемых в каждой выборке необходимо пользоваться Таблицей критических значений критерия χ^2, поскольку критерий Крускала-Уоллиса асимптотически приближается к распределению χ^2 (Носенко И.А., 1981; J. Greene, M. D'Olivera, 1982).

   Количество степеней свободы при этом определяется по формуле: V=c-1 где с - количество сопоставляемых выборок.
3. При множественном сопоставлении выборок достоверные различия между какой-либо конкретной парой (или парами) их могут оказаться стертыми. Это ограничение можно преодолеть, если провести все возможные попарные сопоставления, число которых будет равняться $\frac{1}{2} \cdot [c \cdot (c-1)]$*[6] таких попарных сопоставлений используется, естественно, критерий для *двух* выборок, например

[6] с - количество выборок.

U или φ*.

Пример

В эксперименте по исследованию интеллектуальной настойчивости (Е. В. Сидоренко, 1984) 22 испытуемым предъявлялись сначала разрешимые четырехбуквенные, пятибуквенные и шестибуквенные анаграммы, а затем неразрешимые анаграммы, время работы над которыми не ограничивалось. Эксперимент проводился индивидуально с каждым испытуемым. Использовалось 4 комплекта анаграмм. У исследователя возникло впечатление, что над некоторыми неразрешимыми анаграммами испытуемые продолжали работать дольше, чем над другими, и, возможно, необходимо будет делать поправку на то, какая именно неразрешимая анаграмма предъявлялась тому или иному испытуемому. Показатели длительности попыток в решении неразрешимых анаграмм представлены в Табл. 2.5. Все испытуемые были юношами-студентами технического вуза в возрасте от 20 до 22 лет.

Можно ли утверждать, что длительность попыток решения каждой из 4 неразрешимых анаграмм примерно одинакова?

Таблица 2.5

Показатели длительности попыток решения 4 неразрешимых анаграмм в секундах (7V=22)

	Группа 1: анаграмма ФОЛИТОН (n_1=4)	Группа 2: анаграмма КАМУСТО (n_2=8)	Группа 3: анаграмма СНЕРАКО (n_3=6)	Группа 4: анаграмма ГРУТОСИЛ (n_4=4)
1	145	145	128	60
2	194	210	283	2361
3	731	236	469	2416
4	1200	385	482	3600
5		720	1678	
6		848	2081	
7		905		
8		1080		
Суммы	2270	4549	5121	8437
Средние	568	566	854	2109

Сформулируем гипотезы.

H_0: 4 группы испытуемых, получившие разные неразрешимые анаграммы, не различаются по длительности попыток их решения.

H_1: 4 группы испытуемых, получившие разные неразрешимые анаграммы, различаются по длительности попыток их решения.

Теперь познакомимся с алгоритмом расчетов.

АЛГОРИТМ 5 Подсчет критерия H Крускала-Уоллиса

1. Перенести все показатели испытуемых на индивидуальные карточки.
2. Пометить карточки испытуемых группы 1 определенным цветом, например красным, карточки испытуемых группы 2 - синим, карточки испытуемых групп 3 и 4 - соответственно, зеленым и желтым цветом и т. д. (Можно использовать, естественно, и любые другие обозначения.)
3. Разложить все карточки в единый ряд по степени нарастания признака, не считаясь с тем, к какой группе относятся карточки, как если бы мы работали с одной объединенной выборкой.
4. Проранжировать значения на карточках, приписывая меньшему значению меньший ранг. Надписать на каждой карточке ее ранг. Общее количество рангов будет равняться количеству испытуемых в объединенной выборке.
5. Вновь разложить карточки по группам, ориентируясь на цветные или другие принятые обозначения.
6. Подсчитать суммы рангов отдельно по каждой группе. Проверить совпадение общей суммы рангов с расчетной.

7. Подсчитать значение критерия Н по формуле:

$$H = \left[\frac{12}{N(N+1)} \cdot \Sigma \frac{T_j^2}{n_j}\right] - 3(N+1)$$

где N - общее количество испытуемых в объединенной выборке;

n - количество испытуемых в каждой группе;

T - суммы рангов по каждой группе.

8а. При количестве групп с=3, $n_1 \bullet n_2 \bullet n_3 \leq 5$ определить критические значения и соответствующий им уровень значимости по Табл. IV Приложения 1.

Если $H_{эмп}$ равен или превышает критическое значение $H_{0,05}$, H_0 отвергается.

8б. При количестве групп с>3 или количестве испытуемых $n_1 \bullet n_2 \bullet n_3 > 5$, определить критические значения χ^2 по Табл. IX Приложения 1.

Если $H_{эмп}$ равен или превышает критическое значение χ^2, H_0 отвергается.

Воспользуемся этим алгоритмом при решении задачи о неразрешимых анаграммах. Результаты работы по 1-6 шагам алгоритма представлены в Табл. 2.6.

Таблица 2.6

Подсчет ранговых сумм по группам испытуемых, работавших над четырьмя неразрешимыми анаграммами

Группа 1: анаграмма ФОЛИТОЫ (n_1=4)		Группа 2: анаграмма КАМУСТО (n_2=8)		Группа 3: анаграмма СНЕРАКО (n_3=6)		Группа 4: анаграмма ГРУТОСИЛ (n_4=4)	
Длительность	Ранг	Длительность	Ранг	Длительность	Ранг	Длительность	Ранг
						60	1
				128	2		
145	3.5	145	3.5				
194	5						
		210	6				
		236	7				
				283	8		
		385	9				
				469	10		
				482	11		
		720	12				
731	13						
		848	14				
		905	15				
		1080	16				
1200	17						
				1678	18		
				2081	19		
						2361	20
						2416	21
						3600	22
Суммы	38,5		82,5		68		64
Средние	9,6		10,3		11,3		16,0

Общая сумма рангов =38,5+82,5+68+64=253. Расчетная сумма рангов:

$$\Sigma R_i = \frac{22 \cdot (22+1)}{2} = 253$$

Равенство реальной и расчетной сумм соблюдено.

Теперь определяем эмпирическое значение Н:

$$H_{эмп} = \left[\frac{12}{N \cdot (N+1)} \cdot \sum \frac{T^2}{n} \right] - 3 \cdot (N+1)$$

$$H_{эмп} = \left[\frac{12}{22 \cdot (22+1)} \cdot \left(\frac{38{,}5^2}{4} + \frac{82{,}5^2}{8} + \frac{68^2}{6} + \frac{64^2}{4} \right) \right] - 3 \cdot (22+1) = 2{,}48$$

Поскольку таблицы критических значений критерия Н предусмотрены только для количества групп с = 3, а в данном случае у нас 4 группы, придется сопоставлять полученное эмпирическое значение Н с критическими значениями *у}*. Для этого вначале определим количество степеней свободы V для с=4:

v=c- 1 = 4 - 1 = 3

Теперь определим критические значения по Табл. IX Приложения 1 для *v*=3:

$$\chi^2_{кр.} = \begin{cases} 7.815 \ (\rho \le 0{,}05) \\ 11.345 \ (\rho \le 0{,}01) \end{cases}$$

$H_{эмп.}=2{,}48$

$H_{эмп.}<\chi^2_{кр}$

Ответ: H_0 принимается: 4 группы испытуемых, получившие разные неразрешимые анаграммы, не различаются по длительности попыток их решения.

2.5. S - критерий тенденций Джонкира

Описание этого критерия дается с использованием руководства J.Greene, M.D'Olivera (1982). Он описан также у М. Холлендера, Д.А. Вулфа (1983).

Назначение критерия S

Критерий S предназначен для выявления *тенденций* изменения признака при переходе от выборки к выборке при сопоставлении *трех и более* выборок.

Описание критерия S

Критерий S позволяет нам упорядочить обследованные выборки по какому-либо признаку, например, по креативности, фрустрационной толерантности, гибкости и т.п.

Мы сможем утверждать, что на первом месте по выраженности исследуемого признака стоит выборка, скажем, Б, на втором - А, на третьем - В и т.д. Интерпретация полученных результатов будет зависеть от того, по какому принципу были образованы исследуемые выборки. Здесь возможны два принципиально отличных варианта.

1) Если обследованы выборки, различающиеся по *качественным* признакам (профессии, национальности, месту работы и т. п.), то с помощью критерия S мы сможем упорядочить выборки по *количественно* измеряемому признаку (креативности, фрустрационной толерантности, гибкости и т.п.).

2) Если обследованы выборки, различающиеся или специально сгруппированные по *количественному* признаку (возрасту, стажу работы, социометрическому статусу и др.), то, упорядочивая их теперь уже по другому *количественному* признаку, мы фактически устанавливаем *меру связи* между двумя количественными признаками. Например, мы можем показать с помощью критерия S, что при переходе от младшей возрастной группы к старшей фрустрационная толерантность возрастает, а гибкость, наоборот, снижается.

Меру связи между количественно измеренными переменными можно установить с помощью вычисления коэффициента ранговой корреляции или линейной корреляции (см. Главу 6). Однако критерий тенденций S имеет следующие преимущества перед коэффициентами корреляции:

а) критерий тенденций S более прост в подсчете;

б) он применим и в тех случаях, когда один из признаков варьирует в узком диапазоне, например, принимает всего 3 или 4 значения, в то время как при подсчете ранговой корреляции в этом случае мы получаем огрубленный результат, нуждающийся в

поправке на одинаковые ранги.

Критерий S основан на способе расчета, близком к принципу критерия Q Розенбаума. Все выборки располагаются в порядке возрастания исследуемого признака, при этом выборку, в которой значения в общем ниже, мы помещаем слева, выборку, в которой значения выше, правее, и так далее в порядке возрастания значений. Таким образом, все выборки выстраиваются слева направо в порядке возрастания значений исследуемого признака.

При упорядочивании выборок мы можем опираться на средние значения в каждой выборке или даже на суммы всех значений в каждой выборке, потому что в каждой выборке должно быть *одинаковое 1* количество значений. В противном случае критерий S неприменим j (подробнее об этом см. в разделе "Ограничения критерия S").

Для каждого индивидуального значения подсчитывается ко-\личество значений справа, превышающих его по величине. Если тенденция возрастания признака слева направо существенна, то большая [часть значений справа должна быть выше. Критерий S позволяет определить, преобладают ли справа более высокие значения или нет. Статистика S отражает степень этого преобладания. Чем выше эмпирическое [значение S, тем тенденция возрастания признака является более существенной.

Следовательно, если $S_{эмп}$ равняется критическому значению или превышает его, нулевая гипотеза может быть отвергнута.

Гипотезы

H_0: Тенденция возрастания значений признака при переходе от выборки к выборке является случайной.

H_1: Тенденция возрастания значений признака при переходе от выборки к выборке не является случайной.

Графическое представление критерия

Фактически критерий S позволяет определить, достаточно ли ве-& лика суммарная зона неперекрещивающихся значений в сопоставляемых (выборках: действительно ли в первом ряду значения в общем ниже, чем 1в последующих, во втором - ниже, чем в оставшихся справа последующих и т. д.

Графически это представлено на Рис. 2.7.

На Рис. 2.7(а) у сопоставляемых рядов значений есть непере* 1крещивающиеся зоны, но их суммарная площадь может оказаться 1 слишком небольшой, чтобы признать тенденцию возрастания признака | существенной.

На рис. 2.7(6) сумма неперекрещивающихся зон, по-видимому, достаточно велика, чтобы тенденция возрастания признака была признана достоверной. Точно определить это мы сможем лишь с помощью критерия S.

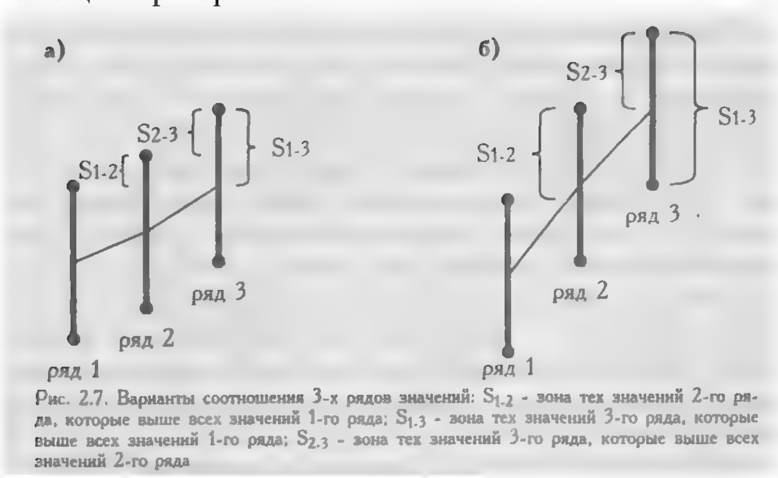


Рис. 2.7. Варианты соотношения 3-х рядов значений: S_{1-2} - зона тех значений 2-го ряда, которые выше всех значений 1-го ряда; S_{1-3} - зона тех значений 3-го ряда, которые выше всех значений 1-го ряда; S_{2-3} - зона тех значений 3-го ряда, которые выше всех значений 2-го ряда

Ограничения критерия S

1. В каждой из сопоставляемых выборок должно быть одинаковое число наблюдений. Если число наблюдений неодинаково, то придется искусственно уравнивать выборки, утрачивая при этом часть полученных наблюдений.

Например, если в двух выборках по 7 наблюдений, а в третьей - И, то 4 из них необходимо отсеять. Для этого карточки с индивидуальными значениями переворачиваются лицевой стороной вниз и перемешиваются, а затем из них случайным образом извлекается 7 карточек. Оставшиеся 4 карточки с индивидуальными значениями не включаются в дальнейшее рассмотрение и в подсчет критерия S. Ясно, что при таком подходе часть информации утрачивается, и общая картина может быть искажена.

Если исследователь хочет избежать этого, ему следует воспользоваться критерием Н, позволяющим выявить различия между тремя и более выборками без указания на направление этих различий (см. параграф 2.4).

2. Нижний порог: не менее 3 выборок и не менее 2 наблюдений в каждой выборке. Верхний порог в существующих таблицах: не более 6 выборок и не более 10 наблюдений в каждой выборке (см. Табл. III Приложения 1 для определения критических значений S). При большем количестве выборок или наблюдений в них придется пользоваться критерием Н Крускала-Уоллиса.

Пример

Выборка претендентов на должность коммерческого директора в Санкт-Петербургском филиале зарубежной фирмы была обследована с помощью Оксфордской методики экспресс-видеодиагностики, использующей диагностические ролевые игры. Были обследованы 20 мужчин в возрасте от 25 до 40 лет, средний возраст 31,5 года. Оценки производились по 15 значимым, с точки зрения зарубежной фирмы, психологическим качествам, обеспечивающим эффективную деятельность на посту коммерческого директора. Одним из этих качеств была "Авторитетность". В конце 8-часового сеанса диагностических ролевых игр и упражнений проводился социометрический опрос участников группы, в котором они должны были ответить на вопрос: "Если бы я сам был представителем фирмы, я выбрал бы на должность коммерческого директора: 1).... 2).... 3)...." Участники знали, что каждый их шаг является материалом для диагностики, и что в данном случае, в частности, проверяется, помимо прочего, их способность к объективному суждению о людях. В результате этой процедуры каждый участник получил то или иное количество выборов от других участников, отражающее его социометрический статус в группе претендентов.

Результаты исследования представлены в Табл. 2.7 (данные Е. В. Сидоренко, И. В. Дермановой, 1991).

Можно ли считать, что группы с разным статусом различаются и по уровню авторитетности, определявшейся независимо от социометрии с помощью экспресс-видеодиагностики?

Таблица 2.7

Показатели по шкале Авторитетности в группах с разным социометрическим статусом (N=20)

Номера испытуемых	Группа 1: 0 выборов (n_1=5)	Группа 2: 1 выбор (n_2=5)	Группа 3: 2-3 выбора (n_3=5)	Группа 4: 4 и более выборов (n_4=5)
1	5	5	5	9
2	5	6	6	9
3	2	7	7	8
4	5	6	7	8
5	4	4	5	7
Суммы	21	28	30	41
Средние	4,2	5,6	6,0	8,2

Сформулируем гипотезы.

H_0: Тенденция повышения значений по шкале Авторитетности при переходе от

группы к группе (слева направо) случайна.

H_1: Тенденция повышения значений по шкале Авторитетности при переходе от группы к группе (слева направо) неслучайна.

Для того, чтобы нам было удобнее подсчитывать количества более высоких значений (S_1), лучше упорядочить значения в каждой группе по их возрастанию (Табл. 2.8).

Таблица 2.8

Расчет критерия S при сопоставлении групп с разным социометрическим статусом по показателю Авторитетности *(N=20)*

Места испытуемых	Группа 1: 0 выборов (n_1=5)		Группа 2: 1 выбор (n_2=5)		Группа 3: 2-3 выбора (n_3=5)		Группа 4[7]: 4 и более выбором (n_4=5)
	Индивидуальные значения	S_i	Индивидуальные значения	S_i	Индивидуальные значения	S_i	Индивидуальные значения
1	2	(15)	4	(10)	5	(5)	7
2	4	(14)	5	(8)	5	(5)	8
3	5	(11)	6	(7)	6	(5)	8
4	5	(11)	6	(7)	7	(4)	9
5	5	(11)	7	(4)	7	(4)	9
Суммы		(62)		(36)		(23)	

После того, как все индивидуальные значения расположены в порядке возрастания, легко подсчитать, сколько значений справа превышают данное значение слева. Начнем с крайнего левого столбца. Значение "2" превышают все 15 значений из трех правых столбцов; значение "4" - 14 значений из трех правых столбцов; значение "5" превышают 11 значений из трех правых столбцов. Полученные количества "превышений" запишем в скобках слева от каждого индивидуального значения, как это сделано в Табл. 2.8.

Расчет для второго столбца производим по тому же принципу. Мы видим, что значение "4" превышают все 10 значений из оставшихся столбцов справа; значение "5" - 8 значений из столбцов справа и т.д.

Сумма всех чисел в скобках (S_i) составит величину А, которую нам нужно будет подставить в формулу для подсчета критерия S. Однако вначале определим максимально возможное значение А, которое мы получили бы, если бы все значения справа были больше значений слева. Эта величина называется величиной В и вычисляется по формуле:

$$B = \frac{c \cdot (c-1)}{2} \cdot n^2$$

где c - количество столбцов (групп);

n - количество испытуемых в каждом столбце (группе). В данном случае:

Эмпирическое значение критерия S вычисляется по формуле:

$$S = 2 \cdot A - B$$

где А - сумма всех "превышений" по всем значениям;

В - максимально возможное количество всех "превышений".

В данном случае:

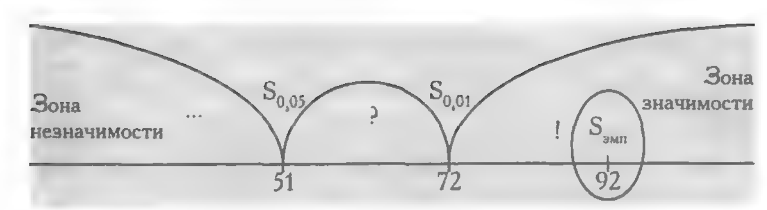
$$S=[2 \cdot [(62+36+23+0)]-150=92$$

По Табл. III Приложения 1 определяем критические значения S для c=4, n=5:

$$S_{кр.} = \begin{cases} 51\ (p \le 0{,}05) \\ 72\ (p \le 0{,}01) \end{cases}$$

Построим "ось значимости".

7 Для крайнего правого столбца S_1 не указываются, поскольку они равны нулю.



Мы помним, что критерий S построен на подсчете количества превышающих значений. Чем это количество больше, тем более достоверные различия мы сможем констатировать. Поэтому "зона значимости" простирается вправо, в область более высоких значений, а "зона незначимости" - влево, в область более низких значений.

$S_{эмп}>S_{кр}$. ($P \leq 0{,}01$)

Ответ: H_0 отвергается. Принимается H_1. Тенденция повышения значений по шкале Авторитетности при переходе от группы к группе не случайна ($p<0{,}01$).

Отвечая на вопрос задачи, мы можем сказать, что группы с разным статусом различаются по показателю Авторитетности, определявшемуся независимо от социометрической процедуры. Критерий S позволяет указать на тенденцию этих изменении: с ростом статуса растут и показатели по шкале Авторитетности. Однако мы имеем дело здесь, конечно же, не с причинно-следственными связями, а с сопряженными изменениями двух признаков. Возможно, оба они изменяются под влиянием одних и тех же общих факторов, например, последовательно проявляющейся в поведении привычки к лидерству, внушающей способности или "харизмы".

Теперь мы можем суммировать все сказанное, алгоритмизировав процесс подсчета критерия S.

АЛГОРИТМ 6 Подсчет критерия S Джонкира

1. Перенести все показатели испытуемых на индивидуальные карточки.

2. Если количества испытуемых в группах не совпадают, уравнять группы, ориентируясь на количество наблюдений в меньшей из групп. Например, если в меньшей из групп *n=3,* то из остальных групп необходимо случайным образом извлечь по три карточки, а остальные отсеять.

Если во всех группах одинаковое количество испытуемых (n<10), можно сразу переходить к п. 3.

3. Разложить карточки первой группы в порядке возрастания признака и занести полученный ряд значений в крайний слева столбец таблицы, затем проделать то же самое для второй группы и занести полученный ряд значений во второй слева столбец, и так далее, пока не будут заполнены все столбцы таблицы.

4. Начиная с крайнего левого столбца подсчитать для каждого индивидуального значения количество превышающих его значений во всех столбцах справа (S_i). Полученные суммы записать в скобках рядом с каждым индивидуальным значением.

5. Подсчитать суммы показателей в скобках по столбцам.

6. Подсчитать общую сумму, просуммировав все суммы по столбцам. Эту общую сумму обозначить как А.

7. Подсчитать максимально возможное количество превышающих значений (В), которое мы получили бы, если бы все значения справа были выше значений слева:

$$B=\frac{c\cdot(c-1)}{2}\cdot n^2$$

где c - количество столбцов (сопоставляемых групп);

n - количество наблюдений в каждом столбце (группе).

8. Определить эмпирическое значение S по формуле:

$S=2\cdot A-B$

9. Определить критические значения S по Табл. III Приложения 1 для данного количества групп (c) и количества испытуемых в каждой группе (n).

Если эмпирическое значение S превышает или по крайней мере равняется кри-

тическому значению, H_0 отвергается.

2.6. Задачи для самостоятельной работы

ВНИМАНИЕ! При выборе критерия рекомендуется пользоваться АЛГОРИТМОМ 7.

Задача 1.

В группе слушателей ФПК по педагогике и психологии назрел глухой конфликт между иногородними слушателями и слушателями, проживавшими в Санкт-Петербурге, где и происходили занятия. В курсе психологического практикума по групповой психологии иногородним слушателям было предложено принять на себя роль петербуржцев и участвовать в споре на их стороне. 7 слушателей были протагонистами - активными игроками, перевоплотившимися в петербуржцев, а 7 других суфлировали им, подсказывая реплики и ссылки на те или иные факты. После этого сеанса социодраматической замены ролей участникам был задан вопрос: "Если принять за 100% психологическую дистанцию между Вами и петербуржцами до дискуссии, то на сколько процентов она сократилась или увеличилась после дискуссии?"

Результаты представлены в Табл. 2.9. Все показатели имеют отрицательный знак, что свидетельствует о сокращении дистанции (Сидоренко Е. В., 1992). Могут ли эти данные использоваться как подтверждение идеи Д. Л. Морено о том, что принятие на себя роли оппонента способствует сближению с ним (Moreno G. L., 1934)?

Таблица 2.9

Показатели сокращения психологической дистанции (в %) после социодраматической замены ролей в группе протагонистов *(n_1=1)ч* суфлеров (n_2=7)

№ испытуемых	Группа 1: протагонисты *(n_1=7)*	Группа 2: суфлеры (n_2=7)
1	75	10
2	30	10
3	25	15
4	10	20
5	30	30
6	20	25
7	50	5

Задача 2.

В исследовании С.К. Скаковского (1990) изучалась проблема психологических барьеров при обращении в службу знакомств у мужчин и женщин. В эксперименте участвовали 17 мужчин и 23 женщины в возрасте от 17 до 45 лет (средний возраст 32,5 года). Испытуемые Должны были отметить на отрезке точку, соответствующую интенсивности внутреннего сопротивления, которое им пришлось преодолеть, чтобы обратиться в службу знакомств. Длина отрезка, отражающая максимально возможное сопротивление, составляла 100 мм. В Табл. 2.10 приведены показатели интенсивности сопротивления, выраженные в миллиметрах.

Можно ли утверждать, что мужчинам приходится преодолевать субъективно более мощное сопротивление?

Таблица 2.10

Показатели интенсивности внутреннего сопротивления при обращении в службу знакомств (в мм)

Группа **1 - мужчины (ni=17)**		Группа 2 - женщины (из—23)	
1	81	1	70
2	80	2	66
3	73	3	66
4	72	4	63
5	72	5	63
6	69	6	61

7	69	7	60
8	65	8	54
9	65	9	47
10	62	10	43
11	60	11	41
12	54	12	40
13	54	13	39
14	43	14	38
15	30	15	38
16	26	16	35
17	26	17	30
		18	27
		19	25
		20	23
		21	17
		22	10
		23	9

Задача 3.

В выборке из 28 мужчин-руководителей подразделений крупного промышленного предприятия Санкт-Петербурга перед началом курса тренинга партнерского общения проводилось обследование с помощью 16-факторного личностного опросника Р. Б. Кеттелла (форма А). В Табл. 2.11 приведены индивидуальные значения испытуемых по фактору N, отражающему житейскую искушенность и проницательность.

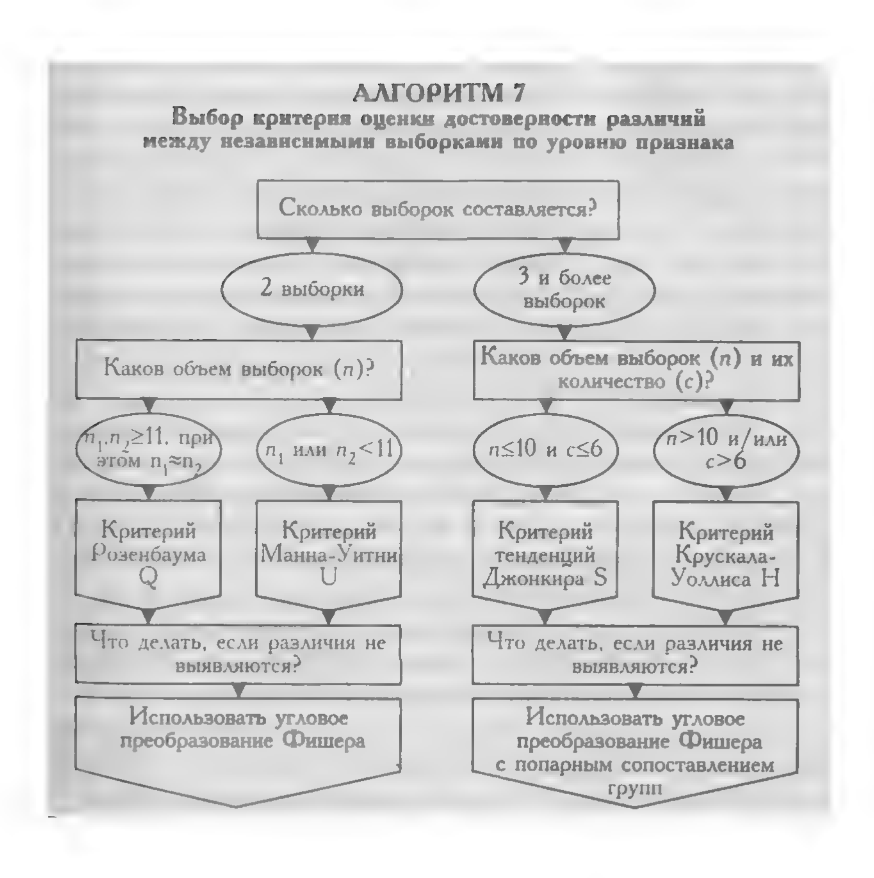
Данные представлены в "сырых" баллах и сгруппированы по четырем возрастным группам. Можно ли утверждать, что есть определенная тенденция изменения значений фактора N при переходе от группы к группе?

Таблица 2.11

Индивидуальное значение по фактору N 16PF в 4 возрастных группах руководителей (по данным Е. В. Сидоренко, 1987)

№ испы-туемых	Группа 1: 26-31 год (n_1=7)	Группа 2: 32-37 лет *(n_2=TL_*	Группа 3: 38-42 года (n_3=7)	Группа 4: 46-52 года (n_4=7)
1	2	11	8	И
2	10	7	12	12
3	5	8	14	9
4	8	12	9	9
5	10	12	16	10
6	7	12	14	14
7	12	9	10	13
Суммы	54	71	83	78
Средние	7,71	10,14	11,86	11,14

2.7. Алгоритм принятия решения о выборе критерия для сопоставлений


АЛГОРИТМ 7
Выбор критерия оценки достоверности различий между независимыми выборками по уровню признака
Сколько выборок составляется?
2 выборки
3 и более выборок
Каков объем выборок (n)?
Каков объем выборок (n) и их количество (c)?
n1,n2≥11, при этом n1≈n2
n1 или n2<11
n≤10 и c≤6
n>10 и/или c>6
Критерий Розенбаума Q
Критерий Манна-Уитни U
Критерий тенденций Джонкира S
Критерий Крускала-Уоллиса H
Что делать, если различия не выявляются?
Что делать, если различия не выявляются?
Использовать угловое преобразование Фишера
Использовать угловое преобразование Фишера с попарным сопоставлением групп

ГЛАВА 3
ОЦЕНКА ДОСТОВЕРНОСТИ СДВИГА В ЗНАЧЕНИЯХ ИССЛЕДУЕМОГО ПРИЗНАКА

3.1. Обоснование задачи исследований изменений

В психологических исследованиях часто бывает важно доказать, что в результате действия каких-либо факторов произошли достоверные изменения ("сдвиги") в измеряемых показателях. К числу таких факторов должен быть отнесен прежде всего фактор времени. Сопоставление показателей, полученных у одних и тех же испытуемых по одним и тем же методикам, но в разное время, дает нам *временной* сдвиг.

Многократные обследования одних и тех же лиц на протяжении достаточно длительного отрезка их жизненного пути, измеряемого иногда десятками лет, представляет собой так называемое лонгитюдинальное исследование, суть которого хорошо известна любому представителю Ленинградской-Петербургской школы психологии. Этот метод позволяет определить генетические связи между фазами психического развития и дать научно обоснованный прогноз дальнейшего психического развития (Ананьев Б.Г., 1976, с. 26-27).

Сопоставление показателей, полученных по одним и тем же методикам, но в разных условиях измерения (например, "покоя" и "стресса"), дает нам *ситуационный* сдвиг. Условия измерения могут изменяться не только реально, но и умозрительно. Например, мы можем попросить испытуемого "представить себе", что он оказался в других условиях измерения: в будущем, в позиции других людей, которые оценивают его как бы со стороны, в состоянии разгневанного отца и т. п. Сопоставляя показатели, измеренные в обычных и воображаемых условиях, мы получаем *умозрительный* сдвиг.

Мы можем создать специальные экспериментальные условия, предположительно влияющие на те или иные показатели, и сопоставить замеры, произведенные до и после экспериментального воздействия. Если сдвиги окажутся статистически достоверными, это позволит нам утверждать, что экспериментальные воздействия были существенными, или эффективными.

Например, мы можем сделать вывод о том, что данная программа тренинга действительно способствует развитию уверенности, или что данный способ внушающего воздействия влияет на изменение отношения испытуемых к той или иной проблеме, или что психодраматическая замена ролей подтверждает постулат Дж.Л. Морено о сближении позиций спорщиков после того, как им пришлось играть роль своего оппонента и т.п.

Во всех этих случаях мы говорим о *сдвиге под влиянием* контролируемых или не контролируемых воздействий. И здесь мы наталкиваемся на методическую трудность, которую оказывается возможным преодолеть только путем введения контрольной группы, которая не испытывала бы на себе воздействия данного экспериментального фактора. Если нет контрольной группы, то сдвиг в экспериментальной группе может объясняться действием самых разных причин: временем суток, в которое производились замеры, важным для испытуемых событием, которое произошло между 1-м и 2-м замерами и по мощности воздействия значительно перекрыло экспериментальный фактор и т. п. Мы никогда не сможем исключить той возможности, что изменения, достигнутые, как нам кажется, в результате наших воздействий, на самом деле объясняются неучтенными причинами. Вот если в экспериментальной группе сдвиги окажутся достоверными, **а** в контрольной группе - недостоверными, то это, действительно, может свидетельствовать об эффективности воздействий. При отсутствии контрольной группы мы констатируем, что сдвиг произошел, но не имеем права приписать его именно данным, изучаемым нами, факторам воздействия.

Допустим, мы установили, что после того, как двум конфликтующим подгруппам пришлось играть роль своих оппонентов в споре, усилилось ощущение понимания этих оппонентов "изнутри" (см. Задачу 1). Но мы не можем исключить возможности, что если

бы мы не проводили психодраматической замены ролей, взаимопонимание все-таки бы улучшилось просто в силу того, что обе подгруппы какое-то время учились и работали вместе.

Бывают случаи, когда мы не располагаем контрольной группой, но зато в нашем распоряжении есть 2 или более экспериментальных групп, различающихся по условиям и способам воздействия на них. Это могут быть, помимо экспериментальных, и разнообразные естественные условия жизни, обучения, работы, общения и даже питания, водоснабжения, географического расположения и т. д. Сопоставление групп, различающихся по этим признакам, позволит нам уточнить' специфическое действие экспериментальных или естественно действующих факторов, хотя при этом нам следует помнить, что воздействие неучтенных факторов может оказаться еще более мощным.

В выводах мы все-таки будет ограничены, если не проверили свои результаты на контрольной труппе, в которой измерения производились параллельно.

Помимо рассмотренных сдвигов: временных, ситуационных, умозрительных и сдвигов под влиянием, - можно рассмотреть еще особую категорию *структурных* сдвигов.

Мы можем сопоставлять между собой разные показатели одних и тех же испытуемых, если они измерены в одних и тех же единицах, по одной и той же шкале. Например, мы можем исследовать перепад между вербальным и невербальным интеллектом, измеренными по методике Д. Векслера, или сопоставлять экспертные оценки эмпатичности и наблюдательности, измеренные по одинаковой 10-балльной шкале, или время решения двух задач, измеренное в секундах, или экзаменационную успешность по разным дисциплинам и т.п.

В принципе, мы могли бы для такого рода "перепадов" использовать критерии оценки достоверности в средних тенденциях для независимых выборок: U - критерий, Q - критерий и угловое преобразование Фишера. Однако, строго говоря, перед нами - зависимые ряды значений, поскольку они измерены на одних и тех же испытуемых, поэтому будет более обоснованным использовать критерии оценки достоверности сдвигов для связанных выборок. Исключение представляют случаи, когда мы сопоставляем величины сдвигов в двух независимых группах испытуемых, например экспериментальной и контрольной (см. Табл. 3.1). Допустим, если мы установили, что положительный сдвиг в сторону улучшения взаимопонимания наблюдается и в экспериментальной, и в контрольной группах, мы можем попробовать доказать, что в экспериментальной группе этот сдвиг достоверно больше, чем в контрольной, и что, следовательно, экспериментальное воздействие все-таки существенно.

Последний важный вопрос касается того, должны ли мы всегда производить оба замера на одной и той же выборке, или "сдвиг" можно изучать на сходных, так называемых "уравновешенных" выборках, совпадающих друг с другом по полу, возрасту, профессии и другим значимым для исследователя характеристикам.

В сущности, допускается сопоставление показателей разных выборок, уравновешенных по всем значимым для исследования признакам. Иными словами, можно уровень тревоги или объем внимания до экзамена измерять у одной подгруппы, а после экзамена - у другой подгруппы, если они "уравновешены". Опыт показывает, однако, что создать "уравновешенные" подгруппы практически невозможно. Мы всегда упираемся в факт существования различий между выделенными подгруппами, которые могут в значительной степени повлиять на результат. В итоге окажется, что мы исследовали не влияние экзаменационного стресса на уровень тревоги или объем внимания, а различия по этому показателю между двумя выделенными подгруппами. К сожалению, в значительной степени это относится и к проблеме сопоставления экспериментальной и контрольной групп: мы почти никогда не можем быть уверены, что выявленные различия объясняются действием исследуемых факторов, а не различиями между двумя выборками.

Многие исследователи обходят эту проблему самым простым образом: они вообще

не заботятся о контрольной группе. Сдвиг есть -значит, воздействие эффективно! И действительно, при отсутствии контрольной выборки тоже можно порассуждать на тему о том, какими же причинами, кроме предполагаемой, могут объясняться полученные сдвиги...

Другой вариант "уравновешивания" - введение параллельных форм теста. В тех случаях, когда на результатах повторных замеров могут сказаться эффекты научения, приходится "до" измерять реакции испытуемого с помощью одного инструмента, а "после" - с помощью другого. В результате на измерениях может отразиться и действие фактора времени, и различия в параллельных формах теста, и непонятно что еще. Создать параллельную форму методики не менее трудно, чем подобрать "уравновешенную" группу испытуемых. И все же, в тех случаях, когда у нас нет другого выхода, приходится прибегать к этому способу.

Суммируем сказанное. В Табл. 3.1 приведена классификация сдвигов и указаны статистические методы, позволяющие оценить их достоверность.

Таблица 3А

Классификация сдвигов и критериев оценки их статистической достоверности

Виды сдвигов	Объект сопоставлений	Условия		Критерии оценки достоверности сдвига
		Количество замеров	Количество групп	
1. Временные, ситуационные, умозрительные, измерительные	Одни и те же показатели, измеренные у одних и тех же испытуемых в разное время, в разных ситуациях в разных представляемых условиях или разными способами	2	1	**G** - критерий знаков; **T** - критерий Вилкоксона
		Зи более	1	**L** - критерий тенденций Пейджа; χ^2_r - критерий Фридмана
2. Сдвига под влиянием экспериментальных воздействий	Одни и те же показатели, измеренные у одних и тех же испытуемых до и после воздействия: а) при отсутствии контрольной группы	2	1	**G** - критерий знаков; **T** - критерий Вилкоксона
		3 и более	1	**L** - критерий тенденций Пейджа; χ^2_r- критерий Фридмана
	б) при наличии контрольной группы	2	2	**Вариант 1** - сопоставление значений "до" и "после" отдельно по экспериментальной и контрольной группам: **G** - критерий знаков; **T** - критерий Вилкоксона **Вариант 2** - сопоставление сдвигов в двух группах: **Q** - критерий; **U** - критерий Манна-

				Уитни; **φ*** - критерий Фишера
		3 и более	2	Сопоставление значений отдельно по экспериментальной и контрольной группам: **L**, - критерий тенденций Пейджа; χ^2, - критерий Фридмана
3. Структурные сдвиги	Разные показатели одних и тех же испытуемых	2	1	**G** - критерий знаков; **T** - критерий Вилкоксона
		3 и более	1	**L** - критерий тенденций Пейджа; χ^2_r - критерий Фридмана

Как следует из Табл. 3.1, при сопоставлении двух, замеров, произведенных на одной и той же (экспериментальной) выборке, применяются критерии знаков G и критерий T Вилкоксона. При сопоставлении трех и более замеров, произведенных на одной и той же выборке, применяются критерий тенденций L Пейджа, а если он неприменим из-за большого объема выборок - критерий χ^2_r Фридмана.

В тех случаях, когда мы хотим оценить различия в интенсивности сдвига в двух группах испытуемых (контрольной и экспериментальной или двух экспериментальных), мы можем использовать различные варианты сопоставлений: 1) производить сопоставления отдельно в двух группах, используя критерии L и χ^2_r; 2) сопоставлять показатели сдвига[8] в двух группах. Поскольку группы независимы, значения сдвигов также независимы, и мы можем применять по отношению к ним уже известные нам критерии Q Розенбаума, U Манна-Уитни и φ* -угловое преобразование Фишера.

3.2. G- критерий знаков

Назначение критерия G

Критерий знаков[9] G предназначен для установления общего *направления* сдвига исследуемого признака.

Он позволяет установить, в какую сторону в выборке в целом изменяются значения признака при переходе от первого измерения ко второму: изменяются ли показатели в сторону улучшения, повышения или усиления или, наоборот, в сторону ухудшения, понижения или ослабления.

Описание критерия G

Критерий знаков применим и к тем сдвигам, которые можно определить лишь *качественно* (например, изменение отрицательного отношения к чему-либо на положительное), так и к тем сдвигам, которые могут быть измерены *количественно* (например, сокращение времени работы над заданием после экспериментального воздействия).

---

[8] Сдвиг - это разность между вторым и первым замерами. Сначала вычисляются разности отдельно для каждой из групп, а уж затем проводятся сопоставления двух рядов разностей (сдвигов), полученных в разных группах. Примером такого сопоставления сдвигов в ощущении психологической дистанции является Задача 1.

[9] Критерий знаков с математической точки зрения является частным случаем биномиального критерий для двух равновероятных альтернатив. При вероятности каждой из альтернатив P=Q=0,50 критерий знаков является зеркальным отражением биномиального критерия (см. параграф 5.3). В некоторых руководствах критерий знаков называют критерием Мак-Немара (McCall R., 1970; Рунион P., 1982).

Во втором случае, однако, если сдвиги варьируют в достаточно широком диапазоне, лучше применять критерий Т Вилкоксона. Он учитывает не только направление, но и интенсивность сдвигов и может оказаться более мощным в определении достоверности сдвигов, чем критерий знаков.

Как правило, исследователь уже в процессе эксперимента может заметить, что у большинства испытуемых показатели во втором замере имеют тенденцию, скажем, повышаться. Однако ему еще требуется доказать, что положительный сдвиг является преобладающим.

Для начала мы назовем сдвиги, которые нам кажутся преобладающими, типичными сдвигами, а сдвиги более редкого, противоположного направления, нетипичными. Если значения показателя повышаются у большего количества испытуемых, то этот сдвиг мы будем считать типичным. Если мы исследуем отношение испытуемых к какому-либо событию или предложению, и после экспериментальных воздействий у большинства испытуемых отрицательное отношение сменилось на положительное, то этот сдвиг мы назовем типичным.

Есть еще, правда, возможность "нулевых" сдвигов, когда реакция не изменяется или показатели не повышаются и не понижаются, а остаются на прежнем уровне. Однако такие "нулевые" сдвиги в критерии знаков исключаются из рассмотрения. При этом количество сопоставляемых пар уменьшается на число таких "нулевых" сдвигов.

Суть критерия знаков состоит в том, что он определяет, не слишком ли много наблюдается "нетипичных сдвигов", чтобы сдвиг в "типичном" направлении считать преобладающим? Ясно, что чем меньше "нетипичных сдвигов", тем более вероятно, что преобладание "типичного" сдвига является преобладающим. $G_{эмп}$ - это количество "нетипичных" сдвигов. Чем меньше $G_{эмп}$, тем более вероятно, что сдвиг в "типичном" направлении статистически достоверен.

Гипотезы

H_0: Преобладание типичного направления сдвига является случайным.

H_1: Преобладание типичного направления сдвига не является случайным.

Графическое представление критерия знаков

На Рис. 3.1 "типичные" сдвиги изображены в виде светлого облака, а нетипичные сдвиги - темного облака. Мы видим, что на рисунке темное облако значительное меньше. Допустим, после выступления оратора большинство слушателей изменили свое отрицательное отношение к какому-то предложению на положительное. Вместе с тем, часть слушателей изменила свое положительное отношение на отрицательное, проявив "нетипичную" реакцию. Критерий знаков позволяет определить, не слишком ли значительная часть слушателей нетипично прореагировала на выступление оратора? Поглощает ли масса светлого облака небольшое темное облако?



Рис. 3.1. Графическое представление положительных и отрицательных сдвигов в форме облаков: светлое облако - положительные сдвиги, темное облако - отрицательные сдвиги

В Таблице V Приложения 1 даны критические значения критерия знаков для разных

n. Поскольку критерий знаков представляет собой одно из трех исключений из общего правила, представим обобщенную "ось значимости" для этого критерия графически (Рис. 3.2)

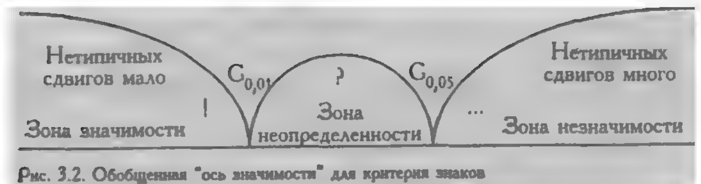


Рис. 3.2. Обобщенная "ось значимости" для критерия знаков

Зона значимости простирается влево, в сторону более низких значений, поскольку чем меньше "нетипичных" знаков, тем достовернее "типичный" сдвиг. Зона незначимости, напротив, простирается вправо, в сторону более высоких значений G. Постепенно "нетипичных" сдвигов становится так много, что теряется само ощущение какого-то преобладания в направленности сдвигов. Зона незначимости характеризует ситуацию, когда сдвиги обоих направлений перемешаны.

Ограничения критерия знаков

Количество наблюдений в обоих замерах - не менее 5 и не более 300.

Пример

В исследовании Г.А. Бадасовой (1994) изучались личностные факторы суггестора, способствующие его внушающему воздействию на аудиторию. В эксперименте участвовало 39 слушателей колледжа и спецфакультета практической психологии Санкт-Петербургского университета 9 мужчин и 30 женщин в возрасте от 18 до 39 лет, средний возраст 23,5 года. Испытуемые выступали в качестве суггерендов, т.е. лиц, по отношению к которым оказывалось внушающее воздействие.

В экспериментальной группе (n_1=16) испытуемые просматривали видеозапись речи суггестора о целесообразности применения физических наказаний в воспитании детей, а в контрольной группе (n_2=23) испытуемые просто читали про себя письменный текст. Содержание речи суггестора и текста полностью совпадали.

До и после предъявления видеозаписи (в экспериментальной группе) и текста (в контрольной группе) испытуемые отвечали на 4 вопроса, оценивая степень согласия с их содержанием по 7-балльной шкале:

1. Я считаю возможным иногда шлепнуть своего ребенка за дело, если
он этого заслужил:

Несогласен 1 2 3 4 5 6 7 Согласен

2. Если, придя домой, я узнаю, что кто-то из близких, бабушка или дедушка, шлепнул моего ребенка за дело, то я буду считать, что это нормально:

Несогласен 1 2 3 4 5 6 7 Согласен

3. Если мне станет известно, что воспитательница детского сада или учительница в школе шлепнула моего ребенка за дело, то я восприму это как должное:

Несогласен 1 2 3 4 5 6 7 Согласен

4. Я бы согласился отдать своего ребенка в школу, где применяется
система физических наказаний по итогам недели:

Несогласен 1 2 3 4 5 6 7 Согласен

Суггестор был подобран по признакам, которые были выявлены в пилотажном исследовании (Бадасова Г. А., 1994). Результаты двух замеров по обеим группам представлены в Табл. 3.2 и Табл. 3 3

Таблица 3.2

Оценки степени согласия с утверждениями о допустимости телесных наказаний до и после предъявления видеозаписи в экспериментальной группе (n_1=16)

№ п/п	Оценки и сдвиги оценок ("после" - "до") по шкалам											
	"Я сам"			"Бабушке"			"Воспитатель"			Школа		
	до	после	сдви	до	посл	сдвиг	до	после	сдвиг	до	после	сдвиг

			г		е							
1	4	4	0	2	4	+2	1	1	0	1	1	0
2	1	1	0	1	1	0	1	1	0	1	1	0
3	5	5	0	4	4	0	4	4	0	1	1	0
4	4	5	+1	3	3	0	2	3	+1	1	2	+1
5	3	3	0	3	4	+1	2	3	+1	1	1	0
6	4	5	+1	5	5	0	1	1	0	1	1	0
7	3	3	0	3	3	0	1	1	0	1	1	0
8	5	6	+1	5	6	+1	3	3	0	2	1	-1
9	6	7	+1	5	7	+2	3	3	0	1	2	+1
10	2	3	+1	2	3	+1	2	1	-1	1	1	0
11	6	6	0	3	3	0	2	1	-1	1	1	0
12	5	5	0	3	5	+2	4	4	0	1	1	0
13	7	7	0	5	5	0	4	4	0	1	1	0
14	5	6	+1	5	6	+1	2	2	0	1	2	+1
15	5	6	+1	5	6	+1	4	3	-1	2	2	0
16	6	7	+1	6	7	+1	4	4	0	2	2	0

Таблица 3.3

Оценки степени согласия с утверждениями о допустимости телесных наказаний до и после предъявления письменного текста в контрольной группе (n_2=23)

№ п/п	Оценки и сдвиги оценок (после - "до") по шкалам											
	"Я сам"			"Бабушка			"Воспитатель"			Школа		
	до	после	сдвиг	ДО	после	сдвиг	до	после	сдвиг	до	после	сдвиг
1	4	4	0	5	5	0	1	1	0	1	1	0
2	7	7	0	7	7	0	7	7	0	4	4	0
3	2	2	0	1	1	0	3	1	-2	1	1	0
4	4	3	-1	3	2	-1	1	1	0	1	1	0
5	3	5	+2	5	5	0	3	3	0	1	1	0
6	2	1	-1	2	1	-1	1	1	0	1	1	0
7	5	5	0	3	3	0	1	1	0	1	1	0
8	2	2	0	2	3	+1	1	3	+2	1	3	+2
9	3	4	+1	3	4	+1	1	1	0	1	6	+5
10	5	5	0	5	5	0	1	1	0	1	1	0
11	5	5	0	1	1	0	1	1	0	1	1	0
12	2	2	0	1	1	0	1	1	0	1	1	0
13	1	1	0	1	1	0	1	2	+1	6	7	+1
14	4	3	-1	7	5	-2	2	4	+2	1	1	0
15	3	4	+1	2	3	+1	1	2	+1	1	1	0
16	4	4	0	3	3	0	1	1	0	1	1	0
17	3	3	0	2	2	0	1	1	0	1	1	0
18	6	6	0	6	6	0	6	6	0	1	3	+2
19	2	2	0	2	1	-1	1	1	0	1	1	0
20	1	2	+1	1	1	0	1	1	0	1	1	0
21	2	2	0	2	2	0	2	1	-1	1	1	0
22	6	6	0	6	6	0	3	3	0	1	1	0
23	3	2	-1	1	2	+1	1	1	0	1	1	0

Вопросы:

1. Можно ли утверждать, что после просмотра видеозаписи о пользе телесных наказаний наблюдается достоверный сдвиг в сторону большего принятия их в экспериментальной группе?

2. Достоверны ли различия по выраженности положительного сдвига между экспериментальной и контрольной группами?

3. Является ли достоверным сдвиг оценок в контрольной группе?

Решение

Подсчитаем сначала количество положительных, отрицательных и нулевых сдвигов по каждой шкале в каждой из выборок. Это необходимо для выявления "типичных" знаков

изменения оценок и значительно облегчит нам дальнейшие расчеты и рассуждения.

Таблица 3.4

Расчет количества положительных, отрицательных и нулевых сдвигов в двух группах суггерендов

Кол-во сдвигов в группах	Шкалы				
	"Я сам"	"Бабушка"	"Воспитатель"	"Школа"	Суммы
1. Экспериментальная группа					
а) положительных	8	9	2	3	22
б) отрицательных	0	0	3	1	4
в) нулевых	8	7	11	12	38
Суммы	16	16	16	16	64
2. Контрольная группа					
а) положительных	4	4	4	4	16
б) отрицательных	4	4	2	0	10
в) нулевых	15	15	17	19	66
Суммы	23	23	23	23	92

Из Табл. 3.4. мы видим, что наиболее типичными являются "нулевые" сдвиги, то есть отсутствие сдвига в оценках после предъявления видеозаписи или письменного текста. И все же, в экспериментальной группе по шкале "Я сам наказываю" и "Бабушка наказывает" положительные сдвиги наблюдаются примерно в половине случаев.

Нам необходимо учитывать только положительные и отрицательные сдвиги, а нулевые отбрасывать. Количество сопоставляемых пар значений при этом уменьшается на количество этих нулевых сдвигов. Теперь для шкалы "Я сам" $n=8$; для шкалы "Бабушка" $n=9$; шкалы "Воспитатель" $n=5$ и шкалы "Школа" $n=4$. Мы видим, что по отношению к последней шкале критерий знаков вообще неприменим, так как количество сопоставляемых пар значений меньше 5.

Мы можем сразу же проверить и гипотезу о преобладании положительного сдвига в ответах по сумме 4 шкал. Сумма положительных и отрицательных сдвигов по 4 шкалам составляет: $n=8+9+5+4=26$.

Сформулируем гипотезы.

H_0: Сдвиг в сторону более снисходительного отношения к телесным наказаниям после внушения является случайным.

H_1: Сдвиг в сторону более снисходительного отношения к телесным наказаниям после внушения является неслучайным.

По Табл. V Приложения 1 определяем критические значения критерия знаков G. Это максимальные количества "нетипичных", менее часто встречающихся, знаков, при которых сдвиг в "типичную" сторону еще можно считать существенным.

1) Шкала "Я сам наказываю"

$n=8$

Типичный сдвиг - положительный.

Отрицательных сдвигов нет.

$$G_{кр} = \begin{cases} 1\,(p \le 0{,}05) \\ 0\,(p \le 0{,}01) \end{cases}$$

$G_{эмп}=0$

$G_{эмп} \le G_{кр}$

H_0 отклоняется. Принимается H_1 ($p<0{,}01$).

2) Шкала "Бабушка наказывает"

$n=9$

Типичный сдвиг - положительный.

Отрицательных сдвигов нет.

$$G_{кр} = \begin{cases} 1 (p \le 0{,}05) \\ 0 (p \le 0{,}01) \end{cases}$$

$G_{эмп}=0$

$G_{эмп} \le G_{кр}$

H_0 отклоняется. Принимается H_1 ($p<0{,}01$).

3) Шкала "Воспитательница наказывает"

n=5

Типичный сдвиг - отрицательный.

Положительных сдвигов - 2.

$G_{кр}=0$ ($p \le 0{,}05$)

$G_{кр}$($p \le 0{,}01$) при данном *n* определить невозможно

$G_{эмп}=2$

$G_{эмп}>G_{кр}$

H_0 принимается.

4) Шкала "Школа наказывает"

n=4

n<5, критерий знаков неприменим.

5) Сумма по 4-м шкалам

n=26

Типичный сдвиг - положительный.

Отрицательных сдвигов - 4

$$G_{кр} = \begin{cases} 8 (p \le 0{,}05) \\ 6 (p \le 0{,}01) \end{cases}$$

$G_{эмп}=4$

$G_{эмп}<G_{кр}$.

H_0 отклоняется. Принимается H_1 ($p<0{,}01$).

Ответ: Сдвиг в сторону более снисходительного отношения к телесным наказаниям в экспериментальной группе после просмотра видеозаписи является неслучайным для шкал "Я сам наказываю", "Бабушка наказывает" и по сумме четырех шкал ($p<0{,}01$ во всех случаях).

Сформулируем гипотезы для контрольной группы.

H_0: Сдвиг в сторону более снисходительного отношения к телесным наказаниям после прочтения текста является случайным.

H_1: Сдвиг в сторону более снисходительного отношения к телесным наказаниям после прочтения текста не является случайным.

Далее действуем по тому же принципу: вначале определяем количество сдвигов в ту или иную сторону (*n*), выявляем типичный сдвиг и количество нетипичных сдвигов ($G_{эмп}$) сопоставляем с критическими значениям G, определяемыми по Табл. V Приложения 1.

1) Шкала "Я сам наказываю"

n=8

Положительных сдвигов - 4, отрицательных сдвигов - 4.

Типичный сдвиг установить невозможно, т.к. положительных и отрицательных сдвигов поровну.

H_0 принимается.

2) Шкала "Бабушка наказывает"

n=8

Положительных сдвигов - 4, отрицательных сдвигов - 4.

H_0 принимается по тем же основаниям, что и для предыдущей шкалы.

3) Шкала "Воспитательница наказывает"

$n=6$

Типичный сдвиг - положительный.

Отрицательных сдвигов - 2.

$G_{кр}=0$ ($p \leq 0,05$)

$G_{кр}$($p \leq 0,01$) при данном n определить невозможно.

$G_{эмп}=2$

$G_{эмп}>G_{кр}$

H_0 принимается.

4) Шкала "Школа наказывает"

Поскольку $n<5$, критерий знаков неприменим.

5) Сумма по 4-м шкалам $n=26$

Типичный сдвиг - положительный. Количество отрицательных сдвигов - 10.

$$G_{кр}=\begin{cases}8\,(p\leq 0.05)\\6\,(p\leq 0.01)\end{cases}$$

$G_{эмп}=10$

$G_{эмп}>G_{кр}$

H_0 принимается.

Ответ: Сдвиг в сторону более снисходительного отношения к телесным наказаниям в контрольной группе является случайным - и по каждой из шкал в отдельности, и по сумме шкал.

Мы можем определенно ответить на 1-ый вопрос задачи: да, можно утверждать, что после просмотра видеозаписи о пользе телесных наказаний наблюдается достоверный сдвиг в пользу большего принятия их в экспериментальной группе. Мы можем ответить и на 3-й вопрос задачи: нет, сдвиг оценок в контрольной группе недостоверен. Однако мы пока не ответили на 2-й вопрос - о том, достоверны ли различия по выраженности положительного сдвига между экспериментальной и контрольной группами?

Дело в том, что нами был избран вариант сопоставлений, предполагающий сравнение значений "после" и "до" экспериментального воздействия отдельно в экспериментальной и контрольной выборках. Для того, чтобы ответить на вопрос 2, необходимо выбрать второй вариант сопоставлений, предусматривающий сравнение сдвигов в двух группах с помощью критериев для сравнения независимых выборок -Q - критерия Розенбаума, U - критерия Манна-Уитни и критерия φ* Фишера (см. Табл. 3.1). Однако такого рода сопоставления, как правило, проводятся только в том случае, если и в экспериментальной, и в контрольной группах выявлен достоверный однонаправленный эффект, и нужно доказать, что в экспериментальной выборке он достоверно больше, выраженнее (см. Задачу 1). В данном же случае нами доказано, что в контрольной выборке не произошло сколько-нибудь значимых изменений, и мы можем этим удовлетвориться.

Казалось бы, мы доказали все, что необходимо: в экспериментальной группе испытуемые стали снисходительнее относиться к телесным наказаниям, а в контрольной группе достоверных сдвигов не обнаружено. Похоже, суггестор, отобранный по выявленным Г. А. Бадасо-вой качествам, действительно повлиял на изменение оценок, и притом именно он, что-то в его личности оказало это воздействие, потому что контрольной группе предъявлялся тот же по содержанию текст, но без суггестора. Однако, на самом деле мы установили лишь то, что в тех случаях, когда наблюдался какой-то сдвиг в оценках, он был скорее положительным, чем отрицательным в экспериментальной группе и скорее случайным в контрольной группе. Все нулевые сдвиги мы отбросили, а ведь они составляют от 43,8 до 50% по тем шкалам, где обнаружен положительный достоверный сдвиг в экспериментальной выборке. Похоже, что многие, очень многие испытуемые экспериментальной выборки просто проигнорировали выступление

суггестора... Однако статистический критерий свидетельствует: положительный сдвиг в оценках достоверен, по крайней мере для первых двух шкал и для тех испытуемых, которые хоть как-то прореагировали на выступление суггестора.

АЛГОРИТМ 8

Расчет критерия знаков G

1. Подсчитать количество нулевых реакций и исключить их из рассмотрения. В результате *n* уменьшится на количество нулевых реакций.

2. Определить преобладающее направление изменений. Считать сдвиги в преобладающем направлении "типичными".

3. Определить количество "нетипичных" сдвигов. Считать это число эмпирическим значением G.

4. По Табл. V Приложения 1 определить критические значения G для данного п.

5. Сопоставить $G_{эмп}$ с $G_{кр}$. Если $G_{эмп}$ меньше $G_{кр}$ или по крайней мере равен ему, сдвиг в типичную сторону может считаться достоверным.

3.3. Т - критерий Вилкоксона

Назначение критерия

Критерий применяется для сопоставления показателей,, измеренных в *двух* разных условиях на *одной* и той же выборке испытуемых.

Он позволяет установить не только *направленность* изменений, но и их *выраженность.* С его помощью мы определяем, является ли сдвиг показателей в каком-то одном направлении более интенсивным, чем в другом.

Описание критерия Т

Этот критерий применим в тех случаях, когда признаки измерены по крайней мере по шкале порядка; и сдвиги между вторым и первым замерами тоже могут быть упорядочены. Для этого они должны варьировать в достаточно широком диапазоне. В принципе, можно применять критерий Т и в тех случаях, когда сдвиги принимают только три значения: —1, 0 и +1, но тогда критерий Т вряд ли добавит что-нибудь новое к тем выводам, которые можно было бы получить с помощью критерия знаков. Вот если сдвиги изменяются, скажем, от —30 до +45, тогда имеет смысл их ранжировать и потом суммировать ранги.

Суть метода состоит в том, что мы сопоставляем выраженность сдвигов в том и ином направлениях по абсолютной величине. Для этого мы сначала ранжируем все абсолютные величины сдвигов, а потом суммируем ранги. Если сдвиги в положительную и в отрицательную сторону происходят случайно, то суммы рангов абсолютных значений их будут примерно равны. Если же интенсивность сдвига в одном из направлений перевешивает, то сумма рангов абсолютных значений сдвигов в противоположную сторону будет значительно ниже, чем это могло бы быть при случайных изменениях.

Первоначально мы исходим из предположения о том, что типичным сдвигом будет сдвиг в более часто встречающемся направлении, а нетипичным, или редким, сдвигом - сдвиг в более редко встречающемся направлении.

Гипотезы

H_0: Интенсивность сдвигов в типичном направлении не превосходит интенсивности сдвигов в нетипичном направлении.

H_1: Интенсивность сдвигов в типичном направлении превышает интенсивность сдвигов в нетипичном направлении.

Графическое представление критерия Т

Сдвиги в противоположные стороны мы можем представить себе в виде двух облаков, как и в критерии знаков. Величина облака зависит не только от количества соответствующих сдвигов, но и от их интенсивности, отраженной в длине стрелок (Рис. 3.3). В сущности, облака: противостоят друг *пруту,* как два воздушных фронта: они не просто соревнуются по величине, они меряются силами! При определенных *n,* а именно при *n*≥18, мы вообще можем отказаться от понятия типичного сдвига. Сдвигов в ту и другую сторону может оказаться поровну, но если 9 меньших сдвигов будут относиться к одному направлению, а 9 больших сдвигов - к противоположному, то мы можем констатировать достоверное преобладание этого противоположного направления сдвигов. Вспомним, что критерий знаков в этом случае не выявил бы никаких достоверных различий.

а) "светлый фронт" преобладает над "темным фронтом" и по количеству сдвигов, и по [illegible]

б) "светлый фронт" преобладает над "темным" только по интенсивности сдвигов, но по количеству сдвигов они равны

в) "светлый фронт" уступает "темному" по количеству сдвигов, но самые интенсивные сдвиги принадлежат "светлому фронту"

Рис. 3.3. Варианты соотношения "светлого" и "темного фронтов" - сдвигов двух разных направленностей

На Рис. З.З(а) "светлый фронт" преобладает над "темным фронтом" и по количеству сдвигов, и по их интенсивности. На Рис. 3.3(6) "светлый фронт* преобладает только по интенсивности сдвигов, но не по их количеству; на Рис. З.З(в) в "светлом фронте" наблюдаются более интенсивные сдвиги, но их меньше, чем в "темном фронте". Здесь критерий знаков мог бы констатировать преобладание изменений, соот-етствующих "темному фронту". Между тем, интенсивность противопо-|жнь1х, хотя и редких, сдвигов, столь велика, что делать какие-то од-ачные выводы было бы опрометчиво.

Ограничения в врнменеанн критерия Т Ввлкоксона

1. Минимальное количество испытуемых, прошедших измерения в двух условиях - 5 человек. Максимальное количество испытуемых - 50 человек, что диктуется верхней границей имеющихся таблиц. Крити-чесхие значения Т приведены в Табл. VI Приложения

2. Нулевые сдвиги из рассмотрения исключаются, и количество наблюдений п уменьшается на количество этих нулевых сдвигов (McCall R., 1970, р. 36). Можно обойти это ограничение, сформулировав гипотезы, включающие отсутствие изменений, например: "Сдвиг в сторону увеличения значений превышает сдвиг в сторону уменьшения значений и тенденцию сохранения их на прежнем уровне".

Пример

В выборке курсантов военного училища (юноши в возрасте от 18 до 20 лет) измерялась способность к удержанию физического волевого усилия на динамометре. Сначала у испытуемых измерялась максимальная мышечная сила каждой из рук, а на следующий день им предлагалось выдерживать, на динамометре с подвижной стрелкой

мышечное усилие, равное 1/2 максимальной мышечной силы данной руки. Почувствовав усталость, испытуемый должен был сообщить об этом экспериментатору, но не прекращать опыт, преодолевая усталость и неприятные ощущения - "бороться, пока воля не иссякнет". Опыт проводился дважды; вначале с обычной инструкцией, а затем, после того, как испытуемый заполнял опросник самооценки волевых качеств по методике А.Ц. Пуни (Пуни А.Ц., 1977), ему предлагалось представить себе, что он уже добился идеала в развитии волевых качеств, и продемонстрировать соответствующее идеалу волевое усилие. Подтвердилась ли гипотеза экспериментатора о том, что обращение к идеалу способствует возрастанию волевого усилия? Данные представлены в Табл. 3.5.

Таблица 3.5

Расчет критерия Т при сопоставлении замеров физического волевого усилия

Код имени испытуемого		Длительность удержания усилия на динамометре (с)		Разность ($f_{после}$- $f_{до}$)	Абсолютное значение разности	Ранговый номер разности
		До измерения волевых качеств и обращения к идеалу ($f_{до}$)	После измерения волевых качеств и обращения к идеалу ($f_{после}$)			
1	Г.	64	25	- 39	39	11
2	Кос.	77	50	- 27	27	8
3	**Крив.**	**74**	**77**	**+ 3**	**5**	**1**
4	Кур.	95	76	- 19	19	6
5	Л.	105	67	- 38	38	9,5
6	М.	83	75	- 8	8	4
7	**Р.**	**73**	**77**	**+ 4**	**4**	**2,5**
8	С.	75	71	- 4	4	2,5
9	Т.	101	*63*	- 38	38	9,5
10	**Х.**	**97**	***122***	**+ 25**	***25***	**7**
11	Ю.	78	60	- 18	18	5
Сумма						66

Для подсчета этого критерия нет необходимости упорядочивать ряды значений по нарастанию признака. Мы можем использовать алфавитный список испытуемых, как в данном случае.

Первый шаг в подсчете критерия Т - вычитание каждого индивидуального значения "до" из значения "после"[10]. Мы видим из Табл. 3.5, что 8 полученных разностей - отрицательные и лишь 3 - положительные. Это означает, что у 8 испытуемых длительность удержания мышечного усилия во втором замере уменьшилась, а у 3 - увеличилась. Мы столкнулись с тем случаем, когда уже сейчас мы не можем сформулировать статистическую гипотезу, соответствующую первоначальному предположению исследователя. Предполагалось, что обращение к идеалу будет увеличивать длительность мышечного усилия, а экспериментальные данные свидетельствуют, что лишь в 3 случаях из 11 этот показатель действительно увеличился. Мы можем сформулировать лишь гипотезу, предполагающую несущественность сдвига этого показателя в сторону снижения.

Сформулируем гипотезы.

H_0: Интенсивность сдвигов в сторону уменьшения длительности мышечного усилия не превышает интенсивности сдвигов в сторону ее увеличения.

H_1: Интенсивность сдвигов в сторону уменьшения длительности мышечного усилия превышает интенсивность сдвигов в сторону ее увеличения.

На следующем шаге все сдвиги, независимо от их знака, должны быть проранжированы по выраженности. В Табл. 3.5 в четвертом слева столбце приведены абсолютные величины сдвигов, а в последнем столбце (справа) - ранги этих абсолютных величин. Меньшему значению соответствует меньший ранг. При этом сумма рангов равна 66, что соответствует расчетной:

[10] Можно вычитать значения "после" из значений "до", это никак не повлияет на расчет критерия. Но лучше во всех случаях придерживаться одной системы, чтобы не запутаться самим.

$$\sum R_i = \frac{N \cdot (N+1)}{2} = \frac{11 \cdot (11+1)}{2} = 66$$

Теперь отметим те сдвиги, которые являются нетипичными, в данном случае - положительными. В Табл. 3.5 эти сдвиги и соответствующие им ранги выделены цветом. Сумма рангов этих "редких" сдвигов и составляет эмпирическое значение критерия Т:

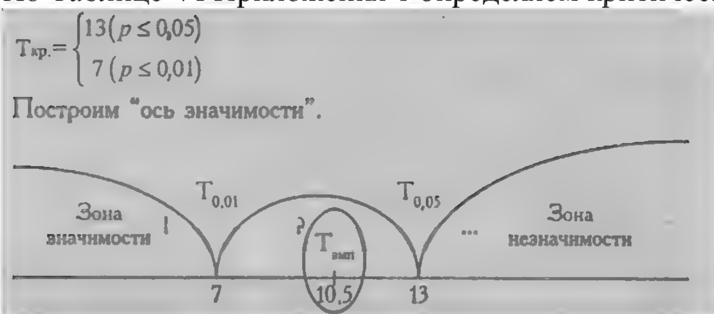
$T = \sum R_r$

где R_r - ранговые значения сдвигов с более редким знаком.

Итак, в данном случае,

$T_{эмп} = 1 + 2{,}5 + 7 = 10{,}5$

По Таблице VI Приложения 1 определяем критические значения Т для $n=11$:



$$T_{кр.} = \begin{cases} 13 \ (p \leq 0{,}05) \\ 7 \ (p \leq 0{,}01) \end{cases}$$

Построим "ось значимости".

Зона значимости в данном случае простирается влево. Действительно, если бы "редких", в данном случае положительных, сдвигов не было совсем, то и сумма их рангов равнялась бы нулю. В данном же случае эмпирическое значение Т попадает в зону неопределенности:

$T_{эмп} < T_{кр}$ (0,05)

Ответ: H_0 отвергается. Интенсивность отрицательного сдвига показателя физического волевого усилия превышает интенсивность положительного сдвига ($p<0{,}05$).

Попытаемся графически отобразить интенсивность отрицательных и положительных сдвигов. На Рис. 3.4 слева сдвиги представлены в секундах, а справа - в своих ранговых значениях. Мы видим, что ранжирование несколько уменьшает площади сопоставляемых облаков, или "фронтов".



Рис. 3.4. Графическое представление отрицательных и положительных сдвигов в длительности удержания мышечного усилия; слева - в секундах; справа - в ранговых значениях

Таким образом, исследователю придется признать, что продолжительность удержания мышечного волевого усилия во втором замере снижается, и этот сдвиг неслучаен. Инструкция, ориентирующая испытуемого на соответствие идеалу в развитии воли, оказалась гораздо менее мощным фактором, чем какая-то иная сила - возможно, мышечное утомление, может быть, разочарование в себе или в возможностях данного психологического эксперимента. А может быть, в момент второго замера просто перестает действовать какой-то мощный фактор, который был активен вначале? На все эти вопросы статистические методы не могут ответить, если в схему эксперимента не включена контрольная группа - в данном случае, выборка, уравновешенная с экспериментальной

группой по всем значимым характеристикам (полу, возрасту, профессии, месту обучения), у которой просто измерили бы вторично волевое усилие через такой же промежуток времени, не призывая соответствовать идеалу в развитии воли.

Представим выполненные действия в виде алгоритма:

АЛГОРИТМ 9
Подсчет критерия Т Вилкоксона

1. Составить список испытуемых в любом порядке, например, алфавитном.
2. Вычислить разность между индивидуальными значениями во втором и первом замерах ("после" - "до"). Определить, что будет считаться "типичным" сдвигом и сформулировать соответствующие гипотезы.
3. Перевести разности в абсолютные величины и записать их отдельным столбцом (иначе трудно отвлечься от знака разности).
4. Проранжировать абсолютные величины разностей, начисляя меньшему значению меньший ранг. Проверить совпадение полученной суммы рангов с расчетной.
5. Отметить кружками или другими знаками ранги, соответствующие сдвигам в "нетипичном" направлении.
6. Подсчитать сумму этих рангов по формуле:

$$T=\Sigma R_r,$$

где R_r - ранговые значения сдвигов с более редким знаком.

7. Определить критические значения Т для данного *n* по Табл. VI Приложения 1. Если $T_{эмп}$ меньше или равен $T_{кр}$, сдвиг в "типичную" сторону по интенсивности достоверно преобладает.

3.4. Критерий χ^2_r Фридмана

Назначение критерия

Критерий χ^2_r применяется для сопоставления показателей, измеренных в *трех или более* условиях на *одной* и той же выборке испытуемых.

Критерий позволяет установить, что величины показателей от условия к условию *изменяются,* но при этом не указывает на направление изменений.

Описание критерия

Данный критерий является распространением критерия Т Вилкоксона на большее, чем 2, количество условий измерения. Однако здесь мы ранжируем не абсолютные величины сдвигов, а сами индивидуальные значения, полученные данным испытуемым в 1, 2, 3 и т. д. замерах.

Например, если у испытуемого в первом замере определена скорость прохождения графического лабиринта 54 сек, во втором замере -42 сек, а в третьем замере - 63 сек, то эти показатели получат ранги, соответственно, 2, 1, 3, поскольку меньшему значению, полученному во втором замере, мы начислим ранг 1, среднему значению, полученному в первом замере - ранг 2, а наибольшему значению, полученному в третьем замере - ранг 3.

После того, как все значения будут проранжированы, подсчитыва-ются суммы рангов по столбцам для каждого из произведенных замеров.

Если различия между значениями признака, полученными в разных условиях, случайны, то суммы рангов по разным условиям будут приблизительно равны. Но если значения признака изменяются в разных условиях каким-то закономерным образом, то в одних условиях будут преобладать высокие ранги, а в других - низкие. Суммы рангов будут достоверно различаться между собой. Эмпирическое значение критерия χ^2_r и указывает на то, насколько различаются суммы рангов. Чем больше эмпирическое

значение χ^2_r, тем более существенные расхождения сумм рангов оно отражает.

Если χ^2_r равняется критическому значению или превышает его, различия статистически Достоверны.

Гипотезы

H_0: Между показателями, полученными (измеренными) в разных условиях, существуют лишь случайные различия.

H_1: Между показателями, полученными в разных условиях, существуют неслучайные различия.

Графическое представление критерия

Графически это будет выглядеть как "пучок" ломаных линий с изломами в одних и тех же местах. На Рис. 3.5 представлены графики изменения времени решения анаграмм' в ходе эксперимента по исследованию интеллектуальной настойчивости. Мы видим, что "сырые" значения пяти испытуемых дают довольно-таки "рассыпающийся" пучок, хотя и с заметной тенденцией к излому в одной и той же точке - на анаграмме № 2. На Рис. 3.6 представлены графики, построенные по ранжированным данным того же исследования. Мы видим, что здесь "пучок" собран практически в одну жирную линию, с единственной выбивающейся из него кривой. В сущности, критерий χ^2_r позволяет нам оценить, достаточно ли согласованно изгибается пучок при переходе от условия к условию. χ^2_r тем больше, чем более выраженными являются различия.

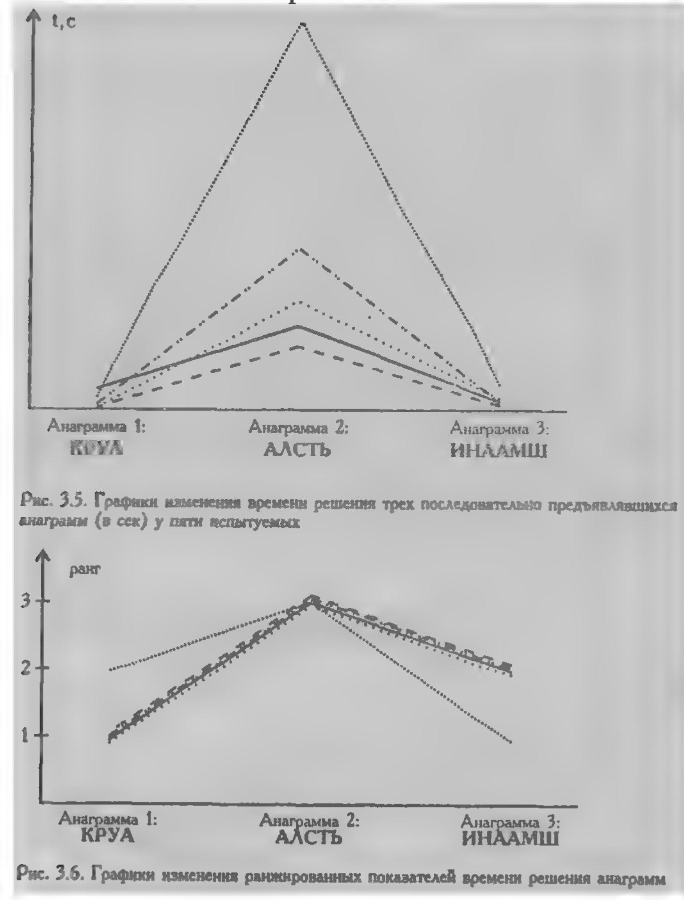


Рис. 3.5. Графики изменения времени решения трех последовательно предъявлявшихся анаграмм (в сек) у пяти испытуемых

Рис. 3.6. Графики изменения ранжированных показателей времени решения анаграмм

Ограничения критерия

1. Нижний порог: не менее 2-х испытуемых ($n \geq 2$), каждый из которых прошел не менее 3-х замеров ($c \geq 3$).

2. При $c=3$, $n \leq 9$, уровень значимости полученного эмпирического значения χ^2_r определяется по Таблице VII-А Приложения 1; при $c=4$, $n \leq 4$, уровень значимости полученного эмпирического значения χ^2_r определяется по Таблице VII-Б Приложения 1; при больших количествах испытуемых или условий полученные эмпирические значения

χ^2_r сопоставляются с критическими значениями χ^2_r, определяемыми по Таблице IX Приложения 1. Это объясняется тем, что χ^2_r имеет распределение, сходное с распределением χ^2_r. Число степеней свободы v определяется по формуле:

v=c—1,

где с - количество условий измерения (замеров).

Пример

На Рис. 3.5. представлены графики изменения времени решения анаграмм в эксперименте по исследованию интеллектуальной настойчивости (Сидоренко Е. В., 1984). Анаграммы нужно было подобрать таким образом, чтобы постепенно подготовить испытуемого к самой трудной - а фактически неразрешимой - задаче. Иными словами, испытуемый должен был постепенно привыкнуть к тому, что задачи становятся все более и более трудными, и что над каждой последующей анаграммой ему приходится проводить больше времени. Достоверны ли различия во времени решения испытуемыми анаграмм?

Таблица 3.5

Показатели времени решения анаграмм (сек.)

Код имени испытуемого	Анаграмма 1: КРУА (РУКА)	Анаграмма 2: АЛСТЬ (СТАЛЬ)	Анаграмма 3: ИНААМШ (МАШИНА)
1. Л-в	5	235*[11]	7
2. П-о	7	604	20
3. К-в	2	93	5
4. Ю-ч	2	171	8
5. Р-о	35	141	7
Суммы	51	1244	47
Средние	10,2	248,8	9,4

Проранжируем значения, полученные по трем анаграммам каждым испытуемым. Например, испытуемый К-в меньше всего времени провел над анаграммой 1 - следовательно, она получает ранг 1. На втором месте у него стоит анаграмма 3 - она получает ранг 2. Наконец, анаграмма 2 получает ранг 3, потому что она решалась им дольше двух других.

Сумма рангов по каждому испытуемому должна составлять 6.

Расчетная общая сумма рангов в критерии определяется по формуле:

$$\sum R_i = n \cdot \frac{c \cdot (c+1)}{2}$$

где n - количество испытуемых

с - количество условий измерения (замеров).

В данном случае,

$$\sum R_i = 5 \cdot \frac{3 \cdot (3+1)}{2} = 30$$

Таблица 3.6

Показатели времени решения анаграмм 1, 2, 3 и их ранги (n=5)

Код имени испытуемого	Анаграмма 1		Анаграмма 2		Анаграмма 3	
	Время (сек)	Ранг	Время (сек)	Ранг	Время (сек)	Ранг
1. Л-в	5	1	235	3	7	2
2. П-о	7	1	604	3	20	2
3. К-в	2	1	93	3	5	2
4. Ю-ч	2	1	171	3	8	2
5. Р-о	35	2	141	3	7	1
Суммы		6		15		9

Общая сумма рангов составляет: 6+15+9—30, что совпадает с расчетной величиной.

[11] *Испытуемый Л-в так и не смог правильно решить анаграмму 2. |4 Е. В. Сидоренко

Мы помним, что испытуемый Л-в провел 3 минуты и 55 сек над решением второй анаграммы, но так и не решил ее. Поскольку он решал ее дольше остальных двух анаграмм, мы имеем право присвоить ей ранг 3. Ведь назначение трех первых анаграмм - подготовить испытуемого к тому, что над следующей анаграммой ему, возможно, придется думать еще дольше, в то время как сам факт нахождения правильного ответа не так существен.

Сформулируем гипотезы.

H_0: Различия во времени, которое испытуемые проводят над решением трех различных анаграмм, являются случайными.

H_1: Различия во времени, которое испытуемые проводят над решением трех различных анаграмм, не являются случайными.

Теперь нам нужно определить эмпирическое значение χ^2_r по формуле:

$$\chi^2_r = \left[\frac{12}{n \cdot c \cdot (c+1)} \cdot \Sigma\left(T_j^2\right)\right] - 3 \cdot n \cdot (c+1)$$

где с - количество условии;

n - количество испытуемых;

T_i - суммы рангов по каждому из условий.

Определим χ^2_r для данного случая:

$$\chi^2_r = \left[\frac{12}{5 \cdot 3 \cdot (3+1)} \cdot \left(6^2 + 15^2 + 9^2\right)\right] - 3 \cdot 5 \cdot (3+1) = 8{,}4$$

Поскольку в данном примере рассматриваются три задачи, то есть 3 условия, с=3. Количество испытуемых n=5. Это позволяет нам воспользоваться специальной таблицей χ^2_r, а именно Табл. VII-А Приложения 1. Эмпирическое значение χ^2_r=8,4 при с=3, n=5 точно соответствует уровню значимости р=0,0085.

Ответ: Но отклоняется. Принимается H_1. Различия во времени, которое испытуемые проводят над решением трех различных анаграмм, неслучайны (р=0,0085).

Теперь мы можем сформулировать общий алгоритм действий по применению критерия χ^2_r.

АЛГОРИТМ 10

Подсчет критерия χ^2_r Фридмана

1. Проранжировать индивидуальные значения первого испытуемого, полученные им в 1-м, 2-м, 3-м и т. д. замерах.

2. Проделать то же самое по отношению ко всем другим испытуемым.

3. Просуммировать ранги по условиям, в которых осуществлялись замеры. Проверить совпадение общей суммы рангов с расчетной суммой.

4. Определить эмпирическое значение χ^2_r по формуле:

$$\chi^2_r = \left[\frac{12}{n \cdot c \cdot (c+1)} \cdot \Sigma\left(T_j^2\right)\right] - 3 \cdot n \cdot (c+1)$$

где с - количество условии;

n - количество испытуемых;

T_i - суммы рангов по каждому из условий.

5. Определить уровни статистической значимости для χ^2_r

а) при с=3, n<9 - по Табл. VII-А Приложения 1;

б) при с=4, n<4 - по Табл. VII-Б Приложения 1.

6. При большем количестве условий и/или испытуемых - определить количество степеней свободы ν по формуле:

ν=с-1,

где с - количество условий (замеров).

По Табл. IX Приложения 1 определить критические значения критерия χ^2 при

данном числе степеней свободы V.

Если $\chi^2_{г\ эмп}$ равен критическому значению χ^2 или превышает его, различия достоверны.

3.5. L - критерий тенденций Пейджа

Описание критерия L дается с использованием руководства J.Greene, M. D'Olivera (1989).

Назначение L - критерия тенденций

Критерий L Пейджа применяется для сопоставления показателей, измеренных в *трех и более* условиях на *одной* и той же выборке испытуемых.

Критерий позволяет выявить *тенденции* в изменении величин признака при переходе от условия к условию. Его можно рассматривать как продолжение теста Фридмана, поскольку он не только констатирует различия, но и указывает на *направление* изменений.

Описание критерия тенденций L

Критерий позволяет проверить наши предположения об определенной возрастной или ситуативно обусловленной динамике тех или иных признаков. Он позволяет объединить несколько произведенных замеров единой гипотезой о тенденции изменения значений признака при переходе от замера к замеру. Если бы не его ограничения, крите-рий был бы незаменим в "продольных", или лонгитюдинальных, исследованиях.

К сожалению, имеющиеся таблицы критических значений рассчитаны только на небольшую выборку *(n<12)* и ограниченное количество сопоставляемых замеров (с<6).

В случае, если эти ограничения не выполняются, приходится использовать критерий $\chi^2_г$ Фридмана, рассмотренный в предыдущем параграфе.

В критерии L применяется такое же ранжирование условий по каждому испытуемому, как и в критерии $\chi^2_г$. Если испытуемый в первом опыте допустил 17 ошибок, во втором - 12, а в третьем - 5, то 1-й ранг получает третье условие, 2-й ранг - второе, а 3-й ранг - первое условие. После того, как значения всех испытуемых будут проранжиро-ваны, подсчитываются суммы рангов по каждому условию. Затем все условия располагаются в порядке возрастания ранговых сумм: на первом месте слева окажется условие с меньшей ранговой суммой, за ним -условие со следующей по величине ранговой суммой, и т. д., пока справа не окажется условие с самой большой ранговой суммой. Далее мы с помощью специальной формулы подсчета L проверяем, действительно ли значения возрастают слева направо. Эмпирическое значение критерия L отражает степень различия между ранговыми суммами, поэтому чем выше значение L, тем более существенны различия.

Гипотезы

H_0: Увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, случайно.

H_1: Увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, неслучайно.

При формулировке гипотез мы имеем в виду новую нумерацию условий, соответствующую предполагаемым тенденциям.

Графическое представление критерия

Используем для иллюстрации пример с предъявлением анаграмм предположительно возрастающей сложности. Замысел экспериментатора состоял в том, чтобы каждая

последующая задача требовала от испытуемых все более длительных раздумий.

Судя по графику на Рис. 3.6, у большинства испытуемых анаграмма 1 стоит на первом ранговом месте, то есть решается быстрее двух других, анаграмма 3 на 2-м ранговом месте, а анаграмма 2 - на 3-м. По-видимому, их следовало бы предъявлять в иной последовательности: 1, 3, 2. График, отражающий такую гипотетическую последовательность задач, представлен на Рис. 3.7.



Рис. 3.7. Графики изменения показателей времени решения (сек.) анаграмм пятью испытуемыми в новой (гипотетической) последовательности их предъявления

Символом достоверной, отчетливой тенденции в изменении показателей при переходе от условия к условию будет достаточно "собранная" ломаная кривая, устремленная кверху или, наоборот, книзу. Если на Рис. 3.6 характерной чертой всех индивидуальных кривых был крутой излом в одной и той же точке графика, то в данном случае на некоторых отрезках повышение кривой характеризуется большей крутизной, а на других - меньшей крутизной. Очевидно, достоверность тенденций будет обеспечиваться именно отрезками более крутого восхождения, но тест тенденций снисходительно распространит этот эффект и на более пологие отрезки.

На Рис. 3.8 графики представлены уже для ранжированных показателей. Здесь уже все различия в крутизне сглажены. L-тест построен на сопоставлении сумм рангов, а ранжирование неизбежно несколько огрубляет полученные показатели. Опыт показывает, однако, что L-тест является достаточно мощным критерием, хотя и ограниченным по сфере применения из-за отсутствия таблиц критических значений для больших *n*.

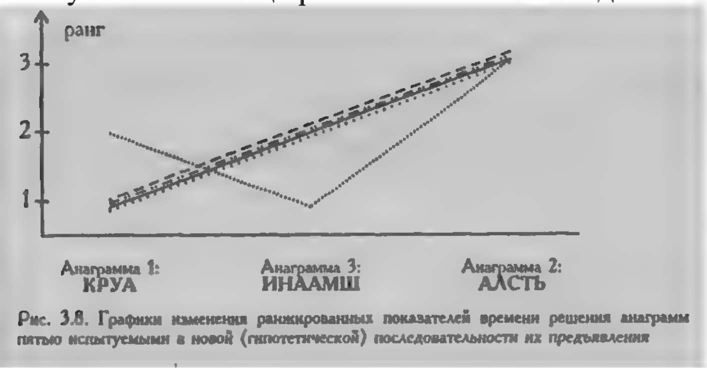


Рис. 3.8. Графики изменения ранжированных показателей времени решения анаграмм пятью испытуемыми в новой (гипотетической) последовательности их предъявления

Ограничения критерия Пейджа

1. Нижний порог - 2 испытуемых, каждый из которых прошел не менее 3-х замеров в разных условиях. Верхний порог - 12 испытуемых и 6 условий ($n \leq 12$, $c \leq 6$). Критические значения критерия L даны по руководству J.Greene, M. D'Olivera (1989). Они предусматривают три уровня статистической значимости: $p \leq 0{,}05$; $p \leq 0{,}01$; $p \leq 0{,}001$.

2. Необходимым условием применения теста является упорядоченность столбцов данных: слева должен располагаться столбец с наименьшей ранговой суммой показателей, справа - с наибольшей. Можно просто пронумеровать заново все столбцы, а потом вести

расчеты не слева направо, а по номерам, но так легче запутаться.

Пример

Продолжим рассмотрение примера с анаграммами. В Табл. 3.7 показатели времени решения анаграмм и их ранги представлены уже в упорядоченной последовательности: анаграмма 1, анаграмма 3, анаграмма 2. Действительно ли время решения увеличивается при такой последовательности предъявления анаграмм?

Таблица 3.7

Показатели времени решения анаграмм 1, 3, 2 и их ранги (n=5)

Код имени испытуемого		Условие 1: Анаграмма 1		Условие *2:* Анаграмма 3		Условие 3: Анаграмма 2	
		Время (сек)	Ранг	Время (сек)	Ранг	Время (сек)	Ранг
1	Л-в	5	1	7	2	235	3
2	П-о	7	1	20	2	604	3
3	К-в	2	1	5	2	93	3
4	Ю-ч	2	1	8	2	171	3
5	Р-о	35	2	7	1	141	3
Суммы		51	6	47	9	1244	15
Средние		10,2		9,4		289	

Сумма рангов составляет: 6+9+5=30. Расчетная сумма:

$$\sum R_i = 5 \cdot \frac{3 \cdot (3+1)}{2} = 30$$

Реально полученная и расчетная суммы совпадают, мы можем двигаться дальше.

Как видно из Табл. 3.7, среднее время решения анаграммы 3 даже меньше, чем анаграммы 1. Однако мы исследуем не среднегруп-повые тенденции, а степень *совпадения индивидуальных тенденций.* Нам важен именно порядок, а не абсолютные показатели времени. Поэтому и формулируемые нами гипотезы - это гипотезы о тенденциях изменения индивидуальных показателей.

Сформулируем гипотезы.

H_0: Тенденция увеличения индивидуальных показателей от первого условия к третьему является случайной.

H_1: Тенденция увеличения индивидуальных показателей от первого условия к третьему не является случайной. Эмпирическое значение L определяется по формуле:

$$L=\Sigma(T_j \cdot j)$$

где T_i - сумма рангов по каждому условию;

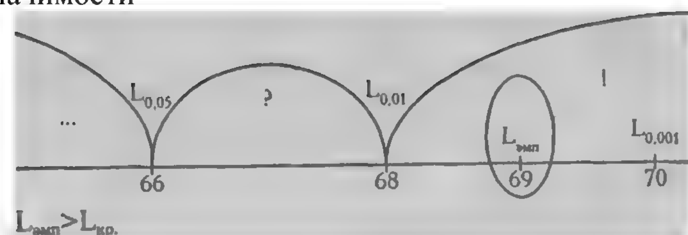
j - порядковый номер, приписанный каждому условию в новой последовательности .

$$L_{эмп}=(6\cdot1)+(9\cdot2)+(15\cdot3)=69$$

По Табл. VIII Приложения 1 определяем критические значения L для данного количества испытуемых: n=5, и данного количества условий: с=3.

$$L_{кр}=\begin{cases} 66\ (p \le 0{,}05) \\ 68\ (p \le 0{,}01) \\ 70\ (p \le 0{,}001) \end{cases}$$

Построим "ось значимости"



Ответ: H_0 отклоняется. Принимается H_1. Тенденция увеличения индивидуальных показателей от первого условия к третьему не является случайной (p<0,01). Последовательность анаграмм: 1(КРУА), 3(ИНААМШ), 2(АЛСТЬ), - будет в большей

степени отвечать замыслу экспериментатора о постепенном возрастании сложности задач, чем первоначально применявшаяся последовательность.

АЛГОРИТМ 11

Подсчет критерия тенденций L Пейджа

1. Проранжировать индивидуальные значения первого испытуемого, полученные им в 1-м, 2-м, 3-ми т. д. замерах.

При этом первым может быть любой испытуемый, например первый по алфавиту имен.

2. Проделать то же самое по отношению ко всем другим испытуемым.

3. Просуммировать ранги по условиям, в которых осуществлялись замеры. Проверить совпадение общей суммы рангов с расчетной суммой.

4. Расположить все условия в порядке возрастания их ранговых сумм в таблице.

5. Определить эмпирическое значение L по формуле:

$L=\Sigma(T_j \cdot j)$

где T_i - сумма рангов по данному условию;

j - порядковый номер, приписанный данному условию в упорядоченной последовательности условий.

6. По Таб.A.VIII Приложения 1 определить критические значения L для данного количества испытуемых п и данного количества условий с. Если $L_{эмп}$ равен критическому значению или превышает его, тенденция достоверна.

3.6. Задачи для самостоятельной работы

ВНИМАНИЕ!
При выборе способа решения задачи рекомендуется пользоваться
АЛГОРИТМОМ 12

Задача 4

В исследовании Г. А. Бадасовой, которое уже рассматривалось как пример к параграфу 3.2, было установлено, что испытуемые по-разному относятся к наказаниям, которые совершают по отношению к их детям разные люди. Например, наказание со стороны самого родителя считается более приемлемым, чем наказание со стороны бабушки, и тем более воспитательницы или учительницы (см. Табл. 3.8).

Таблица 3.8

Оценки степени согласия с утверждениями о допустимости телесных наказаний до предъявления видеозаписи в экспериментальной группе (n=16)

Испытуемые	Условие 1: "Я сам наказываю"	Условие 2: "Бабушка наказывает"	Условие 3: "Учительница наказывает"
1	4	2	1
2	1	1	1
3	5	4	4
4	4	3	2
5	3	3	2
6	4	5	1
7	3	3	1
8	5	5	3
9	6	5	3
10	2	2	2
И	6	3	2
12	5	3	4
13	7	5	4

14	5	5	2
15	5	5	4
16	6	6	4
Суммы	71	60	40

Можно ли говорить о достоверной тенденции в оценках?

Задача 5.

12 участников комплексной программы тренинга партнерского общения, продолжавшегося 7 дней, дважды оценивали у себя уровень владения тремя важнейшими коммуникативными навыками. Первое измерение производилось в первый день тренинга, второе - в последний. Участники должны были также наметить для себя реально достижимый, с их точки зрения, индивидуальный идеал в развитии каждого из навыков. Все измерения производились по 10-балльной шкале. Данные представлены в Табл. 3.9.

Таблица 3.9

Оценки реального и идеального уровней развития коммуникативных навыков (n=12)

Код имени участника		1 измерение						2 измерение					
		Активное слушание		Снижение эмоционального напряжения		Аргументация		Активное слушание		Снижение эмоционально-		Аргументация	
		Реал.	Идеал.	Реал.	Идеал.	Реал.	Идеал.	Реал.	Идеал.	Реал.	Идеал.	Реал.	Идеал.
1	И.	6	9	5	8	5	8	7	10	6	10	7	9
2	Я.	3	5	1	3	4	5	5	7	4	6	5	7
3	Ин.	4	6	4	6	5	8	8	10	7	8	6	8
4	Р.	4	6	4	5	5	7	6	7	5	7	5	7
5	К.	6	9	4	9	4	8	4	10	5	10	5	10
6	Н.	6	8	5	8	3	6	8	9	7	9	6	8
7	Е.	3	8	5	10	2	6	7	8	8	10	5	7
8	Ле.	6	9	5	8	3	7	5	8	7	10	5	9
9	Ли.	6	8	5	9	5	9	7	8	6	9	5	9
10	Т.	5	8	6	9	5	8	7	10	7	10	6	10
11	Ет.	6	8	6	10	3	9	5	10	4	9	3	9
12	Б.	6	8	3	10	4	7	7	9	6	8	5	8

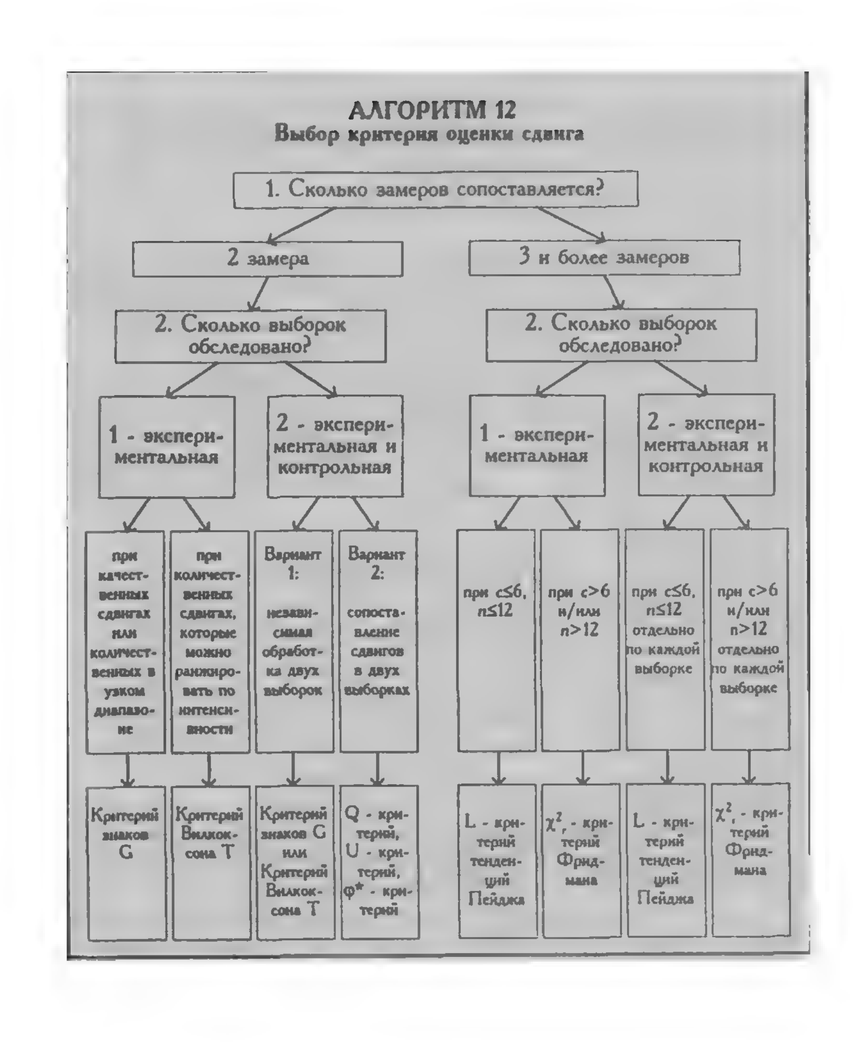
Вопросы:

1. Ощущаются ли участниками достоверные сдвиги в уровне владения каждым из трех навыков после тренинга?

2. Произошли ли по трем группам навыков разные сдвиги, или эти сдвиги для разных навыков примерно одинаковы?

3. Уменьшается ли расхождение между "идеальным" и реальным уровнями владения навыками после тренинга?

3.7. Алгоритм принятия решения о выборе критерия оценки изменений


АЛГОРИТМ 12
Выбор критерия оценки сдвига
1. Сколько замеров сопоставляется?
2 замера
3 и более замеров
2. Сколько выборок обследовано?
2. Сколько выборок обследовано?
1 - экспериментальная
2 - экспериментальная и контрольная
1 - экспериментальная
2 - экспериментальная и контрольная
при качественных сдвигах или количественных в узком диапазоне
при количественных сдвигах, которые можно ранжировать по интенсивности
Вариант 1: независимая обработка двух выборок
Вариант 2: сопоставление сдвигов в двух выборках
при c≤6, n≤12
при c>6 и/или n>12
при c≤6, n≤12 отдельно по каждой выборке
при c>6 и/или n>12 отдельно по каждой выборке
Критерий знаков G
Критерий Вилкоксона T
Критерий знаков G или Критерий Вилкоксона T
Q - критерий, U - критерий, φ* - критерий
L - критерий тенденций Пейджа
χ²r - критерий Фридмана
L - критерий тенденций Пейджа
χ²r - критерий Фридмана

ГЛАВА 4
ВЫЯВЛЕНИЕ РАЗЛИЧИЙ В РАСПРЕДЕЛЕНИИ ПРИЗНАКА

4.1. Обоснование задачи сравнения распределений признака

Распределения могут различаться по средним, дисперсиям, асимметрии, эксцессу и по сочетаниям этих параметров. Рассмотрим несколько примеров.

На Рис. 4.1 представлены два распределения признака. Распределение 1 характеризуется меньшим диапазоном вариативности и меньшей дисперсией, чем распределение 2. В распределении 1 чаще встречаются значения признака, близкие к средней, а в распределении 2 чаще встречаются более высокие и более низкие, чем средняя, значения признака.

Рис. 4.1. Кривые распределения признака с меньшим диапазоном вариативности признака (1) и большим диапазоном распределения признака (2); x - значения признака;
f **- относительная частота их встречаемости**

Именно такое соотношение может наблюдаться в распределении фенотипических признаков у мужчин (кривая 2) и женщин (кривая 1). Фенотипическая дисперсия мужского пола должна быть больше, чем женского (Геодакян В.А., 1974; 1993). Мужчины - это авангардная часть популяции, ответственная за поиск новых форм приспособления, поэтому у них чаще встречаются редкие крайние значения различных фенотипических признаков. Эти отклонения, по мнению В.А. Геодакяна, носят "футуристический" характер, это "пробы", включающие как будущие возможные пути эволюции, так и ошибки (Геодакян В.А., 1974, с. 381). В то же время женская часть популяции ответственна за сохранение уже накопленных изменений, поэтому у них чаще встречаются средние значения фенотипических признаков.

Анализ реально получаемых в исследованиях распределений может позволить нам подтвердить или опровергнуть данные теоретические предположения.

На Рис. 4.2 представлены два распределения, различающиеся по знаку асимметрии: распределение 1 характеризуется положительной асимметрией (левосторонней), а распределение 2 — отрицательной (правосторонней).

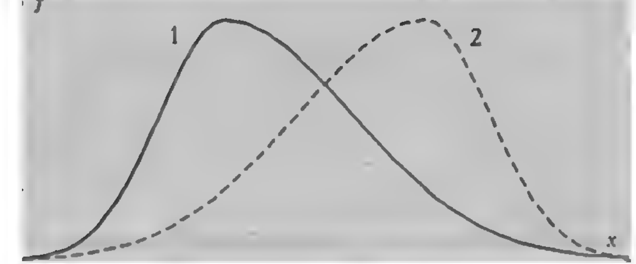


Рис. 4.2. Кривые распределения признака с положительной (левосторонней) асимметрией (1) и отрицательной (правосторонней) асимметрией (2); x - значения признака; (- относительная частота их встречаемости

Данные кривые могут отражать распределение времени решения простой задачи (кривая 1) и трудной задачи (кривая 2). Простую задачу большинство испытуемых

решают быстро, поэтому большая часть значений группируется слева. В то же время сама простота задачи может привести к тому, что некоторые испытуемые будут думать над нею очень, очень долго, дольше даже, чем над сложной. Трудную задачу большинство испытуемых решают в тенденции дольше, чем простую, но в то же время почти всегда находятся люди, которые решают ее мгновенно.

Если мы докажем, что распределения статистически достоверно различаются, это может стать основой для построения классификаций задач и типологий испытуемых. Например, мы можем выявлять испытуемых со стандартным соотношением признаков: простую задачу они решают быстро, а трудную - медленно, — и испытуемых с нестандартным соотношением: простую задачу решают медленно, а трудную - быстро и т.п. Далее мы можем сравнить выявленные группы испытуемых по показателям мотивации достижения, так как известно, что лица с преобладанием стремления к успеху предпочитают задачи средней трудности, где вероятность успеха примерно 0.5, а лица с преобладанием стремления избегать неудачи предпочитают либо очень легкие, либо, наоборот, очень трудные задачи (McClelland D.C, Winter D.G., 1969). Таким образом, и здесь сопоставление форм распределения может дать начало научному поиску.

Часто нам бывает полезно также сопоставить полученное эмпирическое распределение с теоретическим распределением. Например, для того, чтобы доказать, что оно подчиняется или, наоборот, не подчиняется нормальному закону распределения. Это лучше делать с помощью машинных программ обработки данных, особенно при больших объемах выборок. Подробные программы машинной обработки можно найти, например, в руководстве Э.В. Ивантер и А.В. Коросова (1992).

В практических целях эмпирические распределения должны проверяться на "нормальность" в тех случаях, когда мы намерены использовать параметрические методы и критерии. В данном руководстве это относится лишь к методам дисперсионного анализа, поэтому способы проверки совпадения эмпирического распределения с нормальным описаны в Главе 7, посвященной однофакторному дисперсионному анализу.

Традиционные для отечественной математической статистики критерии определения расхождения или согласия распределений - это метод χ^2 К. Пирсона и критерий X Колмогорова-Смирнова.

Оба эти метода требуют тщательной группировки данных и довольно сложных вычислений. Кроме того, возможности этих критериев в полной мере проявляются на больших выборках ($n \geq 30$). Тем не менее они могут оказаться столь незаменимыми, что исследователю придется пренебречь экономией времени и усилий. Например, они незаменимы в следующих двух случаях:

в задачах, требующих доказательства неслучайности предпочтений в выборе из нескольких альтернатив;

в задачах, требующих обнаружения точки максимального расхождения между двумя распределениями, которая затем используется для перегруппировки данных с целью применения критерия φ* (углового преобразования Фишера).

Рассмотрим вначале традиционные методы определения расхождения распределений, а затем возможности использования критерия φ* Фишера.

4,2. χ^2 критерий Пирсона

Назначения критерия

Критерий χ^2 применяется в двух целях;

1) для сопоставления эмпирического распределения признака с *теоретическим* - равномерным, нормальным или каким-то иным;

2) для сопоставления *двух, трех или более эмпирических* распределений одного и

того же признака[12].

Описание критерия

Критерий χ^2 отвечает на вопрос о том, с одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределениях или в двух и более эмпирических распределениях.

Преимущество метода состоит в том, что он позволяет сопоставлять распределения признаков, представленных в любой шкале, начиная от шкалы наименований (см. п. 1.2). В самом простом случае альтернативного распределения "да - нет", "допустил брак - не допустил брака", "решил задачу - не решил задачу" и т. п. мы уже можем применить критерий χ^2.

Допустим, некий наблюдатель фиксирует количество пешеходов, выбравших правую или левую из двух симметричных дорожек на пути из точки А в точку Б (см. Рис. 4.3).

Допустим, в результате 70 наблюдений установлено, что Э\ человек выбрали правую дорожку, и лишь 19 - левую. С помощью критерия χ^2 мы можем определить, отличается ли данное распределение выборов от равномерного распределения, при котором обе дорожки выбирались бы с одинаковой частотой. Это вариант сопоставления полученного *эмпирического* распределения с *теоретическим.* Такая задача может стоять, например, в прикладных психологических исследованиях, связанных с проектированием в архитектуре, системах сообщения и др.

Но представим себе, что наблюдатель решает совершенно другую задачу: он занят проблемами билатерального регулирования. Совпадение полученного распределения с равномерным его интересует гораздо в меньшей степени, чем совпадение или несовпадение его данных с данными других исследователей. Ему известно, что люди с преобладанием правой ноги склонны делать круг против часовой стрелки, а люди с преобладанием левой ноги - круг по ходу часовой стрелки, и что в исследовании коллег[13] преобладание левой ноги было обнаружено у 26 человек из 100 обследованных.

С помощью метода χ^2 он может сопоставить два эмпирических распределения: соотношение 51:19 в собственной выборке и соотношение 74:26 в выборке других исследователей.

Это вариант *сопоставления двух эмпирических* распределений по простейшему альтернативному признаку (конечно, простейшему с математической точки зрения, а отнюдь не психологической).

Аналогичным образом мы можем сопоставлять распределения выборов из трех и более альтернатив. Например, если в выборке из 50 человек 30 выбрали ответ (а), 15 человек - ответ (б) и 5 человек -ответ (в), то мы можем с помощью метода χ^2 проверить, отличается ли это распределение от равномерного распределения или от распределения ответов в другой выборке, где ответ (а) выбрали 10 человек, ответ (б) - 25 человек, ответ (в) - 15 человек.

В тех случаях, если признак измеряется количественно, скажем, **в** баллах, секундах или миллиметрах, нам, быть может, придется объединить все обилие значений признака в несколько разрядов. Например, если время решения задачи варьирует от 10 до 300 секунд, то мы можем ввести 10 или 5 разрядов, в зависимости от объема выборки. Например, это будут разряды: 0-50 секунд; 51-100 секунд; 101-150 секунд, и т. д. Затем мы с помощью метода χ^2 будет сопоставлять частоты встречаемости разных разрядов признака, но в остальном принципиальная схема не меняется.

При сопоставлении эмпирического распределения с теоретическим мы определяем степень расхождения между эмпирическими и теоретическими частотами.

---

[12] На самом деле области применения критерия %^{2} многообразны (см., например: Суходольский Г.В., 1972, с. 295), но в данном руководстве мы ограничиваемся только этими двумя, наиболее часто встречающимися на практике, целями.

[13] Доброхотова Т. А., Брагина Н. Н. Левши. М.: "Книга", 1994.

При сопоставлении двух эмпирических распределений мы определяем степень расхождения между эмпирическими частотами и теоретическими частотами, которые наблюдались бы в случае совпадения двух этих эмпирических распределений. Формулы расчета теоретических частот будут специально даны для каждого варианта сопоставлений.

Чем больше расхождение между двумя сопоставляемыми распределениями, *тем больше* эмпирическое *значение у}.*

Гипотезы

Возможны несколько вариантов гипотез, в зависимости от задач, которые мы перед собой ставим.

Первый вариант:

H_0: Полученное эмпирическое распределение признака не отличается от теоретического (например, равномерного) распределения.

H_1: Полученное эмпирическое распределение признака отличается от теоретического распределения.

Второй вариант:

H_0: Эмпирическое распределение 1 не отличается от эмпирического распределения 2.

H_1: Эмпирическое распределение 1 отличается от эмпирического распределения 2.

Третий вариант:

H_0: Эмпирические распределения 1, 2, 3, ... не различаются между собой.

H_1: Эмпирические распределения 1, 2, 3, ... различаются между собой.

Критерий χ^2 позволяет проверить все три варианта гипотез.

Графическое представление критерия

Проиллюстрируем пример с выбором правой или левой дорожек на пути из точки А в точку Б. На Рис. 4.4 частота выбора левой дорожки представлена левым столбиком, а частота выбора правой дорожки - правым столбиком гистограммы[14]. На оси ординат отмеряются относительные частоты выбора, то есть частоты выбора той или иной дорожки, отнесенные к общему количеству наблюдений. Для левой дорожки относительная частота, которая называется также частостью, составляет 19/70, то есть 0,27, а для правой дорожки 51/70, то есть 0,73.

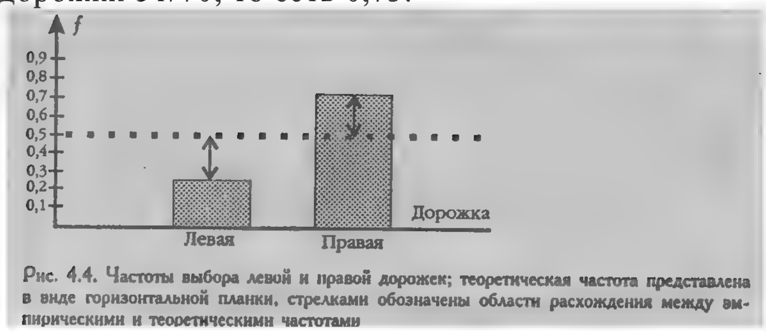


Рис. 4.4. Частоты выбора левой и правой дорожек; теоретическая частота представлена в виде горизонтальной планки, стрелками обозначены области расхождения между эмпирическими и теоретическими частотами

Если бы обе дорожки выбирались равновероятно, то половина испытуемых выбрала бы правую дорожку, а половина - левую. Вероятность выбора каждой из дорожек составляла бы 0,50.

Мы видим, что отклонения эмпирических частот от этой величины довольно значительны. Возможно, различия между эмпирическим и теоретическим распределением окажутся достоверными.

На Рис. 4.5 фактически представлены две гистограммы, но столбики

[14] Гистограмма - это диаграмма, в которой различная величина частот изображается различной высотой столбиков (Плохинский Н. А., 1970, с. 14.)

сгруппированы так, что слева сопоставляются частоты предпочтения левой дорожки в выборе нашего наблюдателя (1) и в выборке Т.А. Доброхотовой и Н.Н. Брагиной (2), а справа - частоты предпочтения правой дорожки в этих же двух выборках.

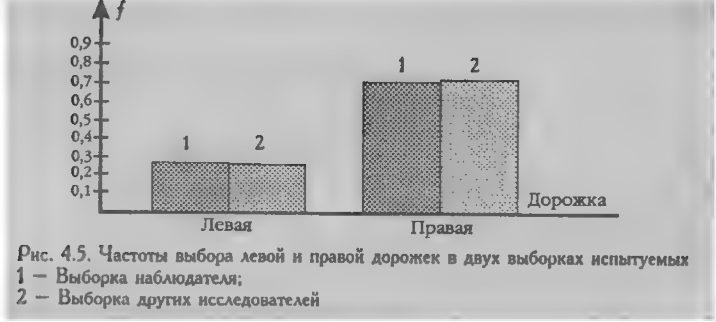


Рис. 4.5. Частоты выбора левой и правой дорожек в двух выборках испытуемых
1 — Выборка наблюдателя;
2 — Выборка других исследователей

Мы видим, что расхождения между выборками очень незначительны. Критерий χ2, скорей всего, подтвердит совпадение двух распределений.

Ограничения критерия

1.Объем выборки должен быть достаточно большим: $n \geqq 30$. При $n<30$ критерий *χ2* дает весьма приближенные значения. Точность критерия повышается при больших *n*.

2. Теоретическая частота для каждой ячейки таблицы не должна быть меньше 5: $f \geq 5$. Это означает, что если число разрядов задано заранее и не может быть изменено, то мы не можем применять метод χ2, не накопив определенного минимального числа наблюдений. Если, например, мы хотим проверить наши предположения о том, что частота обращений в телефонную службу Доверия неравномерно распределяются по 7 дням недели, то нам потребуется 5*7=35 обращений. Таким образом, если количество разрядов (*k)* задано заранее, как в данном случае, минимальное число наблюдений (n_{min}) определяется по формуле: $n_{min}=k*5$.

3. Выбранные разряды должны "вычерпывать" все распределение, то есть охватывать весь диапазон вариативности признаков. При этом группировка на разряды должна быть одинаковой во всех сопоставляемых распределениях.

4. Необходимо вносить "поправку на непрерывность" при сопоставлении распределений признаков, которые принимают всего 2 значения. При внесении поправки значение χ^2 уменьшается (см. Пример с по правкой на непрерывность).

5. Разряды должны быть неперекрещивающимися: если наблюдение отнесено к одному разряду, то оно уже не может быть отнесено ни к какому другому разряду.

Сумма наблюдений по разрядам всегда должна быть равна общему количеству наблюдений.

Правомерен вопрос о том, что считать числом наблюдений - количество выборов, реакций, действий или количество испытуемых, которые совершают выбор, проявляют реакции или производят действия. Если испытуемый проявляет несколько реакций, и все они регистрируются, то количество испытуемых не будет совпадать с количеством реакций. Мы можем просуммировать реакции каждого испытуемого, как, например, это делается в методике Хекхаузена для исследования мотивации достижения или в Тесте фрустрационной толерантности С. Розенцвейга, и сравнивать распределения индивидуальных сумм реакций в нескольких выборках.

В этом случае числом наблюдений будет количество испытуемых. Если же мы подсчитываем частоту реакций определенного типа в целом по выборке, то получаем распределение реакций разного типа, и в этом случае количеством наблюдений будет общее количество зарегистрированных реакций, а не количество испытуемых.

С математической точки зрения правило независимости разрядов соблюдается в обоих случаях: одно наблюдение относится к одному и только одному разряду распределения.

- Можно представить себе и такой вариант исследования, где мы изучаем распределение выборов одного испытуемого. В когнитивно-бихевиоральной терапии, например, клиенту предлагается всякий раз фиксировать точной время появления нежелательной реакции, например, приступов страха, депрессии, вспышек гнева, самоуничижающих мыслей и т. п. В дальнейшем психотерапевт анализирует полученные данные, выявляя часы, в которые неблагоприятные симптомы проявляются чаще, и помогает клиенту

строить индивидуальную программу предупреждения неблагоприятных реакций.

Можно ли с помощью критерия $\chi 2$ доказать, что некоторые часы являются в этом индивидуальном распределении более часто встречающимися, а другие - менее часто встречающимися? Все наблюдения - зависимы, так как они относятся к одному и тому же испытуемому; в то же время все разряды - неперекрещивающиеся, так как один и тот же приступ относится к одному и только одному разряду (в данном случае - часу дня). По-видимому, применение метода $\chi 2$ будет в данном случае некоторым упрощением. Приступы страха, гнева или депрессии могут наступать неоднократно в течение дня, и может оказаться так, что, скажем, ранний утренний, 6-часовой, и поздний вечерний, 12-часовой, приступы обычно появляются вместе, в один и тот же день: в то же время дневной 3-часовой приступ появляется не ранее как через сутки после предыдущего приступа и не менее чем за двое суток до следующего и т. п. По-видимому, речь здесь может идти о сложной математической модели или вообще о чем-то таком, чего нельзя "поверить алгеброй". И тем не менее в практических целях может оказаться полезным использовать критерий для того, чтобы выявить систематическую неравномерность наступления каких-либо значимых событий, выбора, предпочтений и т. п. у одного и того же человека.

Итак, одно и то же наблюдение должно относиться только к одному разряду. Но считать ли наблюдением каждого испытуемого или каждую исследуемую реакцию испытуемого - вопрос, решение которого зависит от целей исследования (см.. напр., Ганзен В.А., Балин В.Д., 1991, с.10).

Главное же "ограничение" критерия χ^2 - то, что он кажется большинству исследователей пугающе сложным.

Попытаемся преодолеть миф о непостижимой трудности критерия χ^2. Чтобы оживить изложение, рассмотрим шутливый литературный пример.



Шутливый пример

В гениальной комедии Н. В. Гоголя "Женитьба" у купеческой дочери Агафьи Тихоновны было пятеро женихов. Одного она сразу исключила из рассмотрения, потому что он был купеческого звания, как и она сама. А из остальных она не знала, кого выбрать: "Уж как трудно решиться, так просто рассказать нельзя, как трудно. Если бы губы Никанора Ивановича да приставить к носу Ивана Кузьмича, да взять сколько-нибудь развязности, какая у Балтазара Балтазарыча, да, пожалуй, прибавить к этому еще дородности Ивана Павловича, я бы тогда тотчас решилась. А теперь поди подумай! просто голова даже стала болеть. Я думаю, лучше всего кинуть жребий" (Гоголь Н.В., 1959, с. 487). И вот Агафья Тихоновна положила бумажки с четырьмя именами в ридикюль, пошарила рукою в ридикюле и вынула вместо одного — всех!

Ей хотелось, чтобы жених совмещал в себе достоинства всех четверых, и, вынимая

все бумажки вместо одной, она бессознательно совершала процедуру выведения средней величины. Но вывести среднюю величину из четверых людей невозможно, и Агафья Тихоновна в смятении. Она влюблена, но не знает, в кого. "Такое несчастное положение девицы, особливо еще влюбленной" (там же, с. 487).

Вся беда в том, что ни Агафья Тихоновна, ни ее тетушка, ни сваха Фекла Ивановна не были знакомы с критерием *χ2!* Именно он мог бы им помочь в решении их проблемы. С его помощью можно было бы попробовать установить, в кого больше влюблена Агафья Тихоновна. Но для этого нам не нужно измерять губы Никанора Ивановича или нос Ивана Кузьмича, или объем талии дородного экзекутора Ивана Павловича; не нужно нам и пускаться на какие-нибудь опасные эксперименты, чтобы определить, насколько далеко простирается развязность Балтазара Балтазарыча. Мы эти их достоинства принимаем как данность потому лишь, что они нравятся Агафье Тихоновне. Мы принимаем их за разряды одного и того же признака, например, направленности взгляда Агафьи Тихоновны: сколько раз она взглянула на губы Никанора Ивановича? На нос Ивана Кузьмича? Благосклонно взирала на дородного Ивана Павловича или развязного Балтазара Бал-тазаровича? Внимательная сваха или тетушка вполне могла бы этот признак наблюдать. Допустим, за полчаса смотрин ею зафиксированы следующие наблюдения.

Агафья Тихоновна:

сидела с опущенными глазами	25 минут;
благосклонно смотрела на Никанора Ивановича	14 раз;
благосклонно смотрела на Ивана Кузьмича	5 раз;
благосклонно смотрела на Ивана Павловича	8 раз;
благосклонно смотрела на Балтазара Балтазарыча	5 раз.[15]

Представим это в виде таблицы.

Таблица 4.1

Распределение взгляда Агафьи Тихоновны между 4 женихами

женихи	Никанор Иванович	Иван Кузьмич	Иван Павлович	Балтазар Балтазарыч	Всего взглядов
Количество взглядов	14	5	8	5	32

Теперь нам нужно сопоставить полученные эмпирические частоты с теоретическими. Если Агафья Тихоновна никому не отдает предпочтения, то данное распределение показателя направленности ее взгляда не будет отличаться от равномерного распределения: она на всех смотрит примерно с одинаковой частотой. Но если достоинства одного из женихов чаще притягивают ее взор, то это может быть основанием для матримониального решения.

Гипотезы

H_0: Распределение взглядов Агафьи Тихоновны между женихами не отличается от равномерного распределения.

H_1: Распределение взглядов Агафьи Тихоновны между женихами отличается от равномерного распределения.

Теперь нам нужно определить теоретическую частоту взгляда при равномерном распределении. Если бы все взгляды невесты распределялась равномерно между 4-мя женихами, то, по-видимому, каждый из них получил бы по $^1/_4$ всех ее взглядов.

Переведем эти рассуждения на более формализованный язык. Теоретическая частота при сопоставлении эмпирического распределения с равномерным определяется по формуле:

---

[15] Все приведенные эмпирические частоты на самом деле пропорциональны количеству благосклонных высказываний невесты о женихах в тексте пьесы.

$$f_{теор} = n/k$$

где *n* - количество наблюдений;

к - количество разрядов признака.

В нашем случае признак - взгляд невесты, направленный на кого-либо из женихов; количество разрядов признака - 4 направления взгляда, по количеству женихов; количество наблюдений - 32.

Итак, в нашем случае:

Теперь мы будем сравнивать с этой теоретической частотой **все** эмпирические частоты.



Рис. 4.6. Сопоставление эмпирических частот взгляда Агафьи Тихоновны на каждого из женихов (столбики гистограммы) с теоретической частотой (горизонтальная планка); темной штриховкой отмечены области расхождений между эмпирическими и теоретическими частотами.

На Рис. 4.6 сопоставления эмпирических частот с теоретической представлены графически. Похоже, что области расхождений достаточно значительны, и Никанор Иванович явно опережает других женихов. Иван Павлович еще может на что-то надеяться, но для Ивана Кузьмича и Балтазара Балтазарыча отставка, по-видимому, неизбежна.

Однако для того, чтобы доказать неравномерность полученного эмпирического распределения, нам необходимо произвести точные расчеты. В методе χ^2 они производятся с точностью до сотых, а иногда и до тысячных долей единицы.

Расчеты будем производить в таблице по алгоритму.

АЛГОРИТМ 13

Расчет критерия $\chi 2$

Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты (первый столбец).

Рядом с каждой эмпирической частотой записать теоретическую частоту (второй столбец).

Подсчитать разности между эмпирической и теоретической частотой по каждому разряду (строке) и записать их в третий столбец.

4. Определить число степеней свободы по формуле:

ν=к-1

где к - количество разрядов признака.

Если ν=1, внести поправку на "непрерывность".

5. Возвести в квадрат полученные разности и занести их в четвертый столбец.

6. Разделить полученные квадраты разностей на теоретическую частоту и записать результаты в пятый столбец.

7. Просуммировать значения пятого столбца. Полученную сумму обозначить как $\chi^2_{эмп.}$

8. Определить по Табл. IX Приложения 1 критические значения для данного числа степеней свободы V.

Если $\chi^2_{эмп}$ меньше критического значения, расхождения между распределениями статистически недостоверны.

Если $\chi^2_{эмп}$ равно критическому значению или превышает его, расхождения между распределениями статистически достоверны.

Все вычисления для данного случая отражены в Табл. 4.2.

Таблица 4.2

Расчет критерия *χ2* при сопоставлении эмпирического распределения взгляда Агафьи Тихоновны между женихами с равномерным распределением

Разряды - женихи		Эмпирическая частота взгляда ($f_{эj}$)	Теоретическая частота ($f_т$)	($f_{эj}$-$f_т$)	$(f_{эj}-f_т)^2$	$(f_{эj}-f_т)^2/ f_т$
1	Никанор Иванович	14	8	+6	35	4.500
2	Иван Кузьмич	5	8	-3	9	1.125
4	Иван Павлович	8	8	0	0	0
5	Балтазар Балтазарыч	5	8	-3	9	1.125
Суммы		32	32	0		6.750

Может показаться, что удобнее суммировать все возведенные в квадрат разности между эмпирическими и теоретическими частотами, а затем уже эту сумму разделить на $f_т$. В данном случае это возможно, так как $f_т$ для всех разрядов одинакова. Однако позже мы увидим, что так бывает далеко не всегда. Нужно быть внимательными или, экономя свое внимание, просто взять за правило всякий раз вычислять $(f_{эi}—f_т)^2/f_т$ до суммирования.

Необходимо также всякий раз убеждаться в том, что сумма разностей между эмпирическими и теоретической частотами (сумма по третьему столбцу) равна 0. Если это равенство не соблюдается, это означает, что в подсчете частот или разностей допущена ошибка. Необходимо найти и устранить ее прежде чем переходить к дальнейшим расчетам.

Алгоритм вычислений, таким образом, выражается формулой:

$$\chi^2 = \sum_{j=1}^{k} \frac{\left(f_{эj} - f_т\right)^2}{f_т}$$

где $f_{эj}$ - эмпирическая частота по j-тому разряду признака; $f_т$ - теоретическая частота; j - порядковый номер разряда; k - количество разрядов признака. В данном случае:

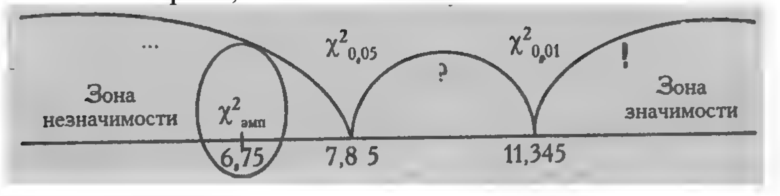
$$\chi^2 = \frac{(14-8)^2}{8} + \frac{(5-8)^2}{8} + \frac{(8-8)^2}{8} + \frac{(5-8)^2}{8} = 6,75$$

Для того, чтобы установить критические значения *%* , нам нужно определить число степеней свободы V по формуле: $\nu=k$-1

где k - количество разрядов. В нашем случае ν=4—1=3. По Табл. IX Приложения 1 определяем:

$$\chi^2_{кр.} = \begin{cases} 7,815 \ (\rho \le 0,05) \\ 11,345 \ (\rho \le 0,01) \end{cases}$$

Построим "ось значимости". Ясно, что чем больше отклонения эмпирических частот от теоретической, тем больше будет величина χ2 . Поэтому зона значимости располагается справа, а зона незначимости -слева.



К сожалению, на основании этих данных тетушка не сможет дать Агафье Тихоновне обоснованного ответа:

$\chi^2_{эмп} < \chi^2_{кр}$.

Ответ: H_0 принимается. Распределение взгляда Агафьи Тихоновны между женихами не отличается от равномерного распределения.

Но, допустим, тетушка на этом не успокоилась. Она стала внимательно следить за тем, сколько раз племянница упомянет в разговоре каждого из женихов. Допустим, ею получено следующее распределение упоминаний Агафьей Тихоновной женихов и их достоинств:

Никанор Иванович - 15 раз,
Иван Кузьмич - 6 раз,
Иван Павлович - 9 раз,
Балтазар Балтазарыч - 6 раз.

Тетушка уже видит, что похоже, Никанор Иванович ("уж такой великатный, а губы, мать моя, - малина, совсем малина") пользуется большей благосклонностью Агафьи Тихоновны, чем все остальные женихи. У нее есть два пути, чтобы это доказать статистически.

1) Суммировать все проявления благосклонности со стороны невесты: взгляды + упоминания в разговоре, - и сопоставить полученное распределение с равномерным. Поскольку количество наблюдений возросло, есть шанс, что различия окажутся достоверными.

2) Сопоставить два эмпирических распределения - взгляда и упоминаний в разговоре, - с тем, чтобы показать, что они совпадают между собой, то есть и во взглядах, и в словах Агафья Тихоновна придерживается одинаковой системы предпочтений. Проанализируем оба варианта сопоставлений. В первом случае мы будем решать уже известную нам задачу сопоставления эмпирического распределения с теоретическим. Во втором случае мы будем сопоставлять два эмпирических распределения. *Первый вариант развития шутливого примера: увеличение количества наблюдений*

Вначале создадим таблицу эмпирических частот, в которой будут суммированы все замеченные проявления благосклонности невесты.

Таблица 4.3

Распределение проявлений благосклонности невесты между женихами

Женихи	Никанор Иванович	Иван Кузьмич	Иван Павлович	Балтазар Балтазарыч	Всего
Количество проявлений	29	11	17	11	68

Теперь сформулируем гипотезы.

H_0: Распределение проявлений благосклонности невесты (взгляды и упоминания в разговоре) не отличается от равномерного распределения. H_1: Распределение проявлений благосклонности невесты отличается от равномерного распределения.. Все расчеты произведем в таблице по алгоритму.

Таблица 4.4

Расчет критерия *χ2* при сопоставлении проявлений благосклонности Агафьи Тихоновны с равномерным распределением

Разряды - женихи		Эмпирические частоты	Теоретическая частота суммарных проявлений	$(f_{эj}-f_т)$	$(f_{эj}-f_т)^2$	$(f_{эj}-f_т)^2/ f_т$
1	Ник. Ив.	29	17	12	144	8,47
2	Ив. Куз.	11	17	-6	36	2,12
4	Ив. Пав.	17	17	0	0	0
5	Бал. Бал.	11	17	-6	36	2,12

Суммы		68	68	0		12,71

$$f_т = {}^{n}/_{k} = 68/4 = 17$$
$$\nu = k-1 = 3$$
$$\chi^2_{кр.} = \begin{cases} 7.815 \ (\rho \le 0.05) \\ 11.345 \ (\rho \le 0.01) \end{cases}$$

$\chi^2_{эмп} = 12,71$

$\chi^2_{эмп} > \chi^2_{кр.}$

Ответ: H_0 отклоняется, принимается H_1. Распределение проявлений благосклонности невесты между женихами отличается от равномерного распределения ($p<0,01$).

На этом примере мы убедились, что увеличение числа наблюдений повышает достоверность результата, если, конечно, в новых наблюдениях воспроизводится прежняя тенденция различий.

Второй вариант развития шутливого примера: сопоставление двух эмпирических распределений

Теперь мы должны ответить на вопрос, одинаковая ли система предпочтений проявляется во взгляде Агафьи Тихоновны и ее словах?

Сформулируем гипотезы. H_0: Распределения невербально и вербально выражаемых предпочтений не различаются между собой.

H_1: Распределения невербально и вербально выражаемых предпочтений различаются между собой.

Для подсчета теоретических частот нам теперь придется составить специальную таблицу (Табл. 4.5). Ячейки в двух столбцах слева обозначим буквами. Для каждой из них теперь будет подсчитана особая, только к данной ячейке относящаяся, теоретическая частота. Это обусловлено тем, что количества взглядов и словесных отзывов невесты о женихах неравны; взглядов 32, а словесных отзывов - 36. Мы должны всякий раз учитывать эту пропорцию.

Таблица 4.5

Эмпирические и теоретические частоты взглядов и упоминаний о жениха

Разряды - женихи		Эмпирические частоты		Суммы	Теоретические частоты	
		взгляда	Упоминаний в разговоре		взгляда	Упоминаний в разговоре
1	Ник. Ив.	14 А	15 Б	29	13,63 А	15,37 Б
2	Ив. Куз.	5 В	6 Г	11	5,17 В	5,83 Г
3	Ив. Пав.	8 Д	9 Е	17	7,99 Д	9,01 Е
4	Бал. Бал.	5 Ж	6 З	11	5,17 Ж	5,83 З
Суммы		32	36	68	32	36

Рассчитаем эту пропорцию. Всего проявлений благосклонности отмечено 68, из них 32 - взгляды и 36 - словесные высказывания. Доля взглядов составит 32/68=0,47; доля упоминаний - 36/68=0,53.

Итак, во всех строках взгляды должны были бы составлять 0,47 всех проявлений по данной строке, а упоминания в разговоре - 0,53 всех проявлений. Теперь, зная суммы проявлений по каждой строке, мы можем рассчитать теоретические частоты для каждой ячейки Табл. 4.5.

$f_{Атеор} = 29*0,47 = 13,63$

$f_{Бтеор} = 29*0,53 = 15,37$

$f_{Втеор} = 11*0,47 = 5,17$

$f_{Гтеор} = 11*0,53 = 5,83$

$f_{Дтеор} = 17*0,47 = 7,99$

$f_{Етеор}$=17*0,53=9,01

$f_{Жтеор}$=110,47=5,17

$f_{Зтеор}$=11*0,53=5,83

Ясно, что сумма теоретических частот по строкам будет равняться сумме всех проявлений по данной строке. Например,

$f_{Атеор}+f_{Бтеор}$=13.63+15,37=29

$f_{Втеор}+f_{Гтеор}$=5,17+5,83=11

$f_{Дтеор}+f_{Етеор}$=7,99+9,01=17 и т.д.

При такого рода подсчетах лучше всякий раз себя проверить. Теперь мы можем вывести общую формулу подсчета $f_{теор}$ для сопоставления двух или более эмпирических распределений:

$$f_{теор} = \frac{\left(\begin{array}{c}\text{Сумма частот по}\\ \text{соответствующей строке}\end{array}\right) \cdot \left(\begin{array}{c}\text{Сумма частот по}\\ \text{соответствующему столбцу}\end{array}\right)}{(\text{Общее количество наблюдений})}$$

Соответствующими строкой и столбцом будут та строка и тот столбец, на пересечении которых находится данная ячейка таблицы. Теперь нам лучше всего сделать развертку Табл. 4.5, представив все ячейки от А до Ж в виде первого столбца - это будет столбец эмпирических частот. Вторым столбцом будут записаны теоретические частоты. Далее будем действовать по уже известному алгоритму. В третьем столбце будет представлены разности эмпирических и теоретических частот, в четвертом - квадраты этих разностей, а в пятом - результаты деления этих квадратов разностей на соответствующие каждой строке теоретические частоты. Сумма в нижнем правом углу таблицы и будет представлять собой эмпирическую величину *%* (Табл. 4.6).

Таблица 4.6

Расчет критерия $\chi 2$ при сопоставлении распределений невербальных и вербальных признаков благосклонности невесты

Ячейки таблицы частот		Эмпирическая частота взгляда ($f_{эj}$)	Теоретическая частота ($f_т$)	($f_{эj}$-$f_т$)	$(f_{эj}-f_т)^2$	$(f_{эj}-f_т)^2/f_т$
1	А	14	13,63	+0,37	0,14	0,01
2	Б	15	15,37	-0,37	0,14	0,01
3	В	5	5,17	-0,17	0,03	0,01
4	Г	6	5,83	+0,17	0,02	0,00
5	Д	8	7,99	+0,01	0,00	0,00
6	Е	9	9,01	-0,01	0,00	0,00
7	Ж	5	5,17	-0,17	0,03	0,01
8	З	6	5,83	+0,17	0,02	0,00
Суммы		68	68	0		0,04

Число степеней свободы при сопоставлении двух эмпирических распределений определяется по формуле:

$\nu=(k-1)\cdot(c-1)$

где k - количество разрядов признака (строк в таблице эмпирических частот);

с - количество сравниваемых распределений (столбцов в таблице эмпирических частот).

В данном случае таблицей эмпирических частот является левая, эмпирическая часть таблицы 4.5, а не на ее развертка (Табл. 4.6). Количество разрядов - это количество женихов, поэтому *k=4.* Количество сопоставляемых распределений с=2. Итак, для данного случая,

ν=(4-1)(2-t)=3

Определяем по Табл. IX Приложения 1 критические значения для ν=3:

$$\chi^2_{кр} = \begin{cases} 7.815\ (\rho \leq 0.05) \\ 11.345\ (\rho \leq 0.01) \end{cases}$$

$\chi^2_{эмп}=0{,}04$

$\chi^2_{эмп}<\chi^2_{кр}$.

Ответ: H_0 принимается. Распределения невербально и вербально выражаемых невестой предпочтений не различаются между собой.

Итак, Агафья Тихоновна весьма последовательна в проявлении своих предпочтений, хотя, по-видимому, сама этого пока не замечает.

Иллюстрация 2



Третий вариант развития шутливого примера: сопоставление встречных выборов

К сожалению, в этом пункте мы от комедии вынуждены перейти к драме - истинной драме любви. Ибо, судя по тексту пьесы, проявляемые женихами признаки влюбленности и симпатии по отношению к невесте отнюдь не соответствуют ее собственной системе предпочтений. У Ивана Павловича, а, главное, у Никанора Ивановича, которому невестой отдается столь явное предпочтение, проскальзывают в разговоре по большей части как раз отрицательные и задумчиво-неодобрительные отзывы о невесте: "Нос велик... Нет, не то, не то... Я даже думаю, что вряд ли она знакома с обхождением высшего общества. Да и знает ли она еще по-французски".

Благосклонных отзывов ("А сказать правду - мне понравилась она потому, что полная женщина" и т. п.) поступило:

от Никанора Ивановича - ни одного;

от Ивана Кузьмича - 15*

от Ивана Павловича - 6*

от Балтазара Балтазарыча - 18.

Попробуем ответить на вопрос: согласуются ли распределения (благосклонных отзывов невесты о женихах и женихов о невесте?

Мы видим, что это действительно особая задача. Мы сопоставляем два эмпирических распределения с совпадающей классификацией разрядов, но в одном случае это распределение реакций одного человека на четверых других, а в другом случае это реакции четырех человек на одного и того же человека.

Такая модель взаимных реакций может использоваться отнюдь не только в области брачных консультаций, но и в решении задач "построения команды", выбора заместителя, подбора пар в тех видах деятельности, где требуется активное постоянное взаимодействие, в исследованиях социальной перцепции и взаимного влияния, в тренинге сенситивности и др.

Сформулируем гипотезы.

H_0: Распределение положительных отзывов невесты совпадает с распределением положительных отзывов женихов.

H_1: Распределение положительных отзывов невесты не совпадает с распределением положительных отзывов женихов.

Построим таблицу для подсчета теоретических частот.

Таблица 4.7

Эмпирические и теоретические частоты положительных высказываний невесты о женихах и женихов о невесте

		Эмпирические частоты		Суммы	Теоретические частоты	
Разряды-женихи		Положительных высказываний невесты о женихах	Положительных высказываний женихов о невесте		Положительных высказываний невесты о женихах	Положительных высказываний женихов о невесте
1	Ник. Ив.	15 А	0 Б	15	7,20 А	7,80 Б
2	Ив. Куз.	6 В	15 Г	21	10,08 В	10,92 Г
3	Ив. Пав.	9 Д	6 Е	15	7,20 Д	7,80 Е
4	Бал. Бал.	6 Ж	18 З	24	11,52 Ж	12,48 З
Суммы		36	39	75	36	39

Теоретические частоты рассчитываем по уже известной формуле:

$$f_{теор} = \frac{\left(\text{Сумма частот по соответствующей строке}\right)\cdot\left(\text{Сумма частот по соответствующему столбцу}\right)}{\left(\text{Общее количество наблюдений}\right)}$$

$f_{А\ теор}$=15*36/75=7,20
$f_{Б\ теор}$=15*39/75=7,80
$f_{В\ теор}$=21*36/75=10,08
$f_{Г\ теор}$=21*39/75=10,92
$f_{Д\ теор}$=15*36/75=7,20
$f_{Е\ теор}$=15*39/75=7,80
$f_{Ж\ теор}$=24*36/75=11,52
$f_{З\ теор}$=24*39/75=12,48

Суммы теоретических частот по строкам совпадают. Все дальнейшие расчеты выполним в таблице по алгоритму.

Таблица 4.8

Расчет критерия $\chi 2$ при сопоставлении распределений высказываний невесты о женихах и женихов о невесте

Ячейки таблицы частот		Эмпирическая частота взгляда ($f_{эj}$)	Теоретическая частота ($f_т$)	($f_{эj}$-$f_т$)	$(f_{эj}-f_т)^2$	$(f_{эj}-f_т)^2/f_т$
1	А	15	7,20	+7,80	60,84	8,45

2	Б	0	7,80	-7,80	60,84	7,80
3	В	6	10,08	-4,08	16,65	1,65
4	Г	15	10,92	+4,08	16,65	1,52
5	Д	9	7,20	+1,80	3,24	0,45
6	Е	6	7,80	-1,80	3,24	0,42
7	Ж	6	11,52	-5,52	30,47	2,64
8	З	18	12,48	+5,52	30,47	2,44
Суммы		75	75	0		25,37

Определим число степеней свободы V по количеству строк *k* и столбцов с в левой части Табл. 4.7: (*k*=4, с=2).

$\nu=(k-1)(c-1)$

Критические значения $\chi2$ для $\nu=3$ нам уже известны:

$$\chi^2_{кр.} = \begin{cases} 7{,}815\ (\rho \le 0{,}05) \\ 11{,}345\ (\rho \le 0{,}01) \end{cases}$$

$$\chi^2_{эмп}=25{,}37$$

$$\chi^2_{эмп} > \chi^2_{кр.}$$

Ответ: H_0 отвергается. Принимается H_1. Распределение положительных отзывов предпочтений невесты не совпадает с распределением положительных отзывов женихов ($\rho<0{,}01$).

Итак, если бы Иван Кузьмич Подколесин не сбежал, Агафью Тихоновну могло бы ожидать не меньшее разочарование: предпочитаемый ею Никанор Иванович, "тонкого поведения человек", ее отвергает.

Мы не рассмотрели лишь третью группу возможных гипотез в методе $\chi2$. Они, как мы помним, касаются сопоставлений одновременно 3 и более распределений. Принцип расчетов там такой же, как и при сопоставлении двух эмпирических распределений. Это касается и формулы расчета теоретических частот, и алгоритма последующих расчетов.

Рассмотрим особые случаи в применении метода $\chi2$.

Особые случаи в применении критерия

1. В случае, если число степеней свободы $\nu=1$, т. е. если признак принимает всего 2 значения, необходимо вносить поправку на непрерывность[16].

2. Если признак варьирует в широком диапазоне (например, от 10 до 140 сек. и т.п.), возникает необходимость укрупнять разряды.

Особый случай 1: поправка на непрерывность для признаков, которые принимают всего 2 значения

Поправка на непрерывность вносится при следующих условиях: а) когда эмпирическое распределение сопоставляется с равномерным распределением, и количество разрядов признака $k=2$, а $\nu=k—1=1$;

б) когда сопоставляются два эмпирических распределения, и количество разрядов признака равно 2, т.е. и количество строк $k=2$, и количество столбцов с=2, и $\nu=(k—1)*(c—1)=1$.

Вариант "а": поправка на непрерывность при сопоставлении эмпирического распределения с равномерным. Это тот случай сопоставлений, когда мы, говоря простым языком, проверяем, поровну ли распределились частоты между двумя значениями признака.

Пример с поправкой на непрерывность.

В исследовании порогов социального атома[17] профессиональных психологов просили определить, с какой частотой встречаются в их записной книжке мужские и

---

[16] Поправка на непрерывность при $\nu=1$ предназначена для корректировки несоответствия между дискретным биномиальным распределением и непрерывным рас пределением (Рунион Р., 1982, с. 39.)

[17] Социальный атом "... состоит из всех отношений между человеком и окружающими его людьми, которые в данный момент тем или иным образом с ним связаны" (Moreno J. L., 1951.)

женские имена коллег-психологов. Попытаемся определить, отличается ли распределение, полученное по записной книжке женщины-психолога X, от равномерного распределения. Эмпирические частоты представлены в Табл. 4.9

Таблица 4.9

Эмпирические частоты встречаемости имен мужчин и женщин в записной книжке психолога X

Мужчин	Женщин	Всего человек
22	45	67

Сформулируем гипотезы.

H_0: Распределение мужских и женских имён в записной книжке X не отличается от равномерного распределения.

H_1: Распределение мужских и женских имен в записной книжке X отличается от равномерного распределения.

Количество наблюдений *n*=67; количество значений признака *k=2*. Рассчитаем теоретическую частоту:

Число степеней свободы *v=k* -1=1.

Далее все расчеты производим по известному алгоритму, но с одним добавлением: перед возведением в квадрат разности частот мы должны уменьшить абсолютную величину этой разности на 0,5 (см. Табл. 4.10, четвертый столбец).

Таблица 4.10

Расчет критерия % при сопоставлении эмпирического распределения имен с теоретическим равномерным распределением

Разряды – принадлежность к тому или иному полу		Эмпирическая частота взгляда ($f_{эj}$)	Теоретическая частота ($f_т$)	($f_{эj}$-$f_т$)	($f_{эj}$-$f_т$-0,5)	($f_{эj}$-$f_т$-0,5)^2	($f_{эj}$-$f_т$-0,5)^2/ $f_т$
1	Мужчины	22	33,5	-11,5	11	121	3,61
2	Женщины	45	33,5	+11,5	11	121	3,61
Суммы		67	67	0			7,22

Для v=l определяем по Табл. IX Приложения 1 критические значения:

$$\chi^2_{кр} = \begin{cases} 3{,}841\ (\rho \le 0{,}05) \\ 6{,}635\ (\rho \le 0{,}01) \end{cases}$$

$\chi^2_{эмп}$=7,22

$\chi^2_{эмп}>\chi^2_{кр}$

Ответ: H_0 отклоняется, принимается H_1. Распределение мужских и женских имен в записной книжке психолога X отличается от равномерного распределения (p<0,01).

Вариант "б": поправка на непрерывность при сопоставлении двух эмпирических распределений

Попытаемся определить, различаются ли распределения мужских и женских имен у психолога X и психолога С, тоже женщины. Эмпирические частоты приведены в Табл. 4.11.

Таблица 4.11

Эмпирические частоты встречаемости имен мужчин и женщин в записных книжках психолога X. и психолога С.

	Мужчин	Женщин	Всего человек
Психолог X.	22 А	45 Б	67
Психолог С.	59 В	109 Г	168
Суммы	81	154	235

Сформулируем гипотезы. H_0: Распределения мужских и женских имен в двух записных книжках

не различаются.

H_1: Распределения мужских и женских имен в двух записных книжках различаются между собой. Теоретические частоты рассчитываем по уже известной формуле:

$$f_{теор} = \frac{(Сумма\ частот\ по\ соотв-щей\ строке)*(Сумма\ частот\ по\ соотв-щему\ столбцу)}{(Общее\ количество\ наблюдений)}$$

А именно, для разных ячеек таблицы эмпирических частот,

$f_{А\ ТЕОР}$=67*81/235=23,09

$f_{Б\ ТЕОР}$ =67*154/235=43.91

$f_{В\ теор}$=168*81/235=57,91

$f_{Г\ теор}$=168*154/235=110,09

Число степеней свободы ν=(*k*—1)*(с—1)=1 Все дальнейшие расчеты проводим по алгоритму (Табл. 4.12)

Таблица 4.12

Расчет критерия при сопоставлении двух эмпирических распределений мужских и женских имен

Ячейки таблицы эмпирических частот		Эмпирическая частота взгляда ($f_{эj}$)	Теоретическая частота ($f_т$)	($f_{эj}$-$f_т$)	($f_{эj}$-$f_т$-0,5)	$(f_{эj}-f_т-0,5)^2$	$(f_{эj}-f_т-0,5)^2/f_т$
1	А	22	23,09	-1,09	0,59	0,35	0,015
2	Б	45	43,91	+1,09	0,59	0,35	0,008
3	В	59	57,91	+1,09	0,59	0,35	0,006
4	Г	109	110,09	-1,09	0,59	0,35	0,003
Суммы		235	235,00	0			0,032

Критические значения $\chi 2$ при ν=1 нам известны по предыдущему примеру:

$$\chi^2_{кр.} = \begin{cases} 3{,}841\ (\rho \le 0{,}05) \\ 6{,}635\ (\rho \le 0{,}01) \end{cases}$$

$\chi^2_{эмп}$=0,03

$\chi^2_{эмп} < \chi^2_{кр.}$

Ответ: H_0 принимается. Распределения мужских и женских имен в записных книжка двух психологов совпадают.

Поправки на непрерывность и всех остальных подсчетов можно избежать, если использовать по отношению к подобного рода задачам метод φ* Фишера (см. параграф 5.4).

Особый случай 2: укрупнение разрядов признака, который варьирует в широком диапазоне значений

Если признак варьирует в широком диапазоне значений, например, от 10 до 140 сек или от 0 до 100 мм и т. п., то вряд ли мы сможем принимать каждое значение признака за самостоятельный разряд:

10 сек, И сек, 12 сек и т. д. до 100 сек. Одно из ограничений критерия χ^2 состоит в том, что теоретически на каждый разряд должно приходиться не менее 5 наблюдений: $f_{теор} \ge 5$. Если у признака 90 значений, и каждое из них принимается за самостоятельный разряд, то необходимо иметь не менее 5*90=450 наблюдений! Если же наблюдений меньше 450, то придется укрупнять разряды до тех пор, пока на каждый разряд не будет приходиться по 5 наблюдений. Это не означает, что в ка-ждом разряде *реально* должно быть 5 наблюдений; это означает, что теоретически на каждый разряд их приходится по 5. Рассмотрим это на примере.

Пример с укрупнением разрядов признака

Тест Мюнстерберга для измерения избирательности перцептивного внимания в адаптированном варианте М.Д. Дворяшиной (1976) предъявлялся студентам факультета психологии Ленинградского университета (n_1=156) и артистам балета Мариинского театра (n_2=85). Материал методики состоит из бланка с набором букв русского алфавита, в случайном порядке перемежающихся. Среди этого фона скрыто 24 слова разной степени сложности: "факт", "хоккей", "любовь", "конкурс", "психиатрия" и т.п. Задача испытуемого возможно быстрее отыскать их и подчеркнуть (Дворяшина М.Д., 1976, с. 124). Совпадают ли распределения количества ошибок (пропусков слов) в двух выборках (Табл. 4.13)?

Таблица 4.13

Эмпирические частоты пропуска слов в тесте Мюнстерберга в двух выборках испытуемых (по данным М.Д. Дворяшиной, Е.В. Сидоренко, 1973)

разряды	Эмпирические частоты пропуска слов		
	В группе студентов (n_1=156)	В группе артистов балета (n_2=85)	Суммы
I. 0 пропусков	93	22	115
II. 1 пропуск	27	20	47
III. 2 пропуска	11	16	27
IV. 3 пропуска	15	4	19
V. 4 пропуска	5	3	8
VI. 5 пропусков	3	11	14
VII. 6 пропусков	2	3	5
VIII. 7 пропусков	0	3	3
IX. 8 пропусков	0	2	2
X. 9 пропусков	0	1	1
Суммы	156	85	241

Сформулируем гипотезы.

H_0: Распределения ошибок (пропусков слов) в выборках студентов и артистов балета не различаются между собой.

H_1: Распределения ошибок (пропусков слов) в выборках студентов и артистов балета различаются между собой.

Прежде чем перейти к расчету теоретических частот, обратим внимание на последние 4 значения признака, от 6 пропусков и ниже. Очевидно, что $f_{теор}$ для любой из ячеек последних 4 строк таблицы будет меньше 5. Например, для ячейки, отмеченной кружком:

$$f_{теор} = \frac{(Сумма\ частот\ по\ соотв-ей\ строке) * (Сумма\ частот\ по\ соотв-ему\ столбцу)}{(Общее\ количество\ наблюдений)}$$

$f_{теор}$=5*85/241=1,763

Полученная теоретическая частота меньше 5.

Для того, чтобы решить, какие разряды нам следует укрупнить, чтобы $f_{теор}$ была не меньше 5, выведем формулу расчета минимальной суммы частот по строке по формуле:

$$Минимальная\ сумма\ по\ строке = \frac{f_{теор\ минимальная} * (общее\ количество\ наблюдений)}{сумма\ частот\ по\ столбцу\ с\ наименьшим\ n}$$

В данном случае столбцом с наименьшим количеством наблюдений является столбец, относящийся к выборке артистов балета (n=85). Определим минимальную сумму частот для каждой строки: Минимальная сумма по строке =5*241/85=14,16 Мы видим, что для получения такой суммы нам недостаточно объединения последних 4 строк Табл. 4.13, так как сумма частот по ним меньше 14 (5+3+2+1=11), а нам

необходима сумма частот, превышающая 14. Следовательно, придется объединять в один разряд пять нижних строк Табл. 4.13: теперь любое количество пропусков от 5 до 9 будет составлять один разряд.

Однако это еще не все. Мы видим, далее, что в строке "4 пропуска" сумма составляет всего 8. Значит, ее необходимо объединить со следующей строкой. Теперь и 3, и 4 пропуска будут входить в один разряд. Все остальные суммы по строкам больше 14, поэтому мы не нуждаемся в дальнейшем укрупнении разрядов.

Эмпирические частоты по укрупненным разрядам представлены в Табл. 4.14.

Таблица 4.14

Эмпирические частоты пропуска слов по укрупненным разрядам в двух выборках испытуемых

Разряды	Эмпирические частоты пропуска слов				
	В группе студентов (n_1=156)		В группе артистов балета (n_2=85)		Суммы
I. 0 пропусков	93	А	22	Б	115
II. 1 пропуск	27	В	20	Г	47
III. 2 пропуска	11	Д	16	Е	27
IV. 3-4 пропуска	20	Ж	7	З	27
V. 5-9 пропусков	5	И	20	К	25
Суммы	156		85		241

Исследователю бывает огорчительно терять информацию, заведомо утрачиваемую при укрупнении разрядов. Например, в данном случае нас может интересовать, удалось ли сохранить специфический для второй выборки спад частот на 3 и 4 пропусках и резкий их подъем на 5 пропусках (Рис. 4.7).

Сравним графики на Рис. 4.7 и Рис. 4.8. Мы видим, что спад частот во второй выборке на 3-х и 4-х пропусках сохранился, а спад на 2-х пропусках в первой выборке стал еще более заметным. В то же время все возможные различия в частотах в диапазоне от 5-и до 9-и пропусков теперь оцениваются только глобально, по соотношению общих сумм частот в этих диапазонах. По графику на Рис. 4.8 мы уже не можем определить, какое максимальное количество пропусков встречается в первой группе и какое - во второй. Сопоставление распределений на этом конце становится более грубым.

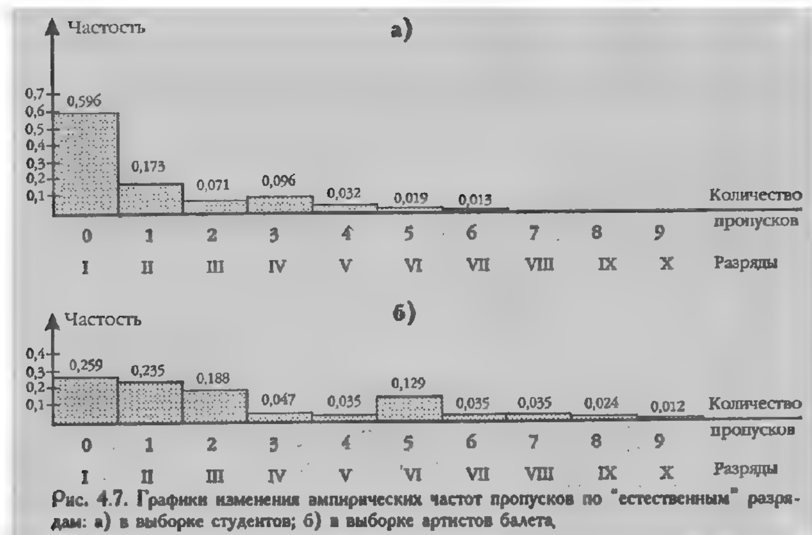


Рис. 4.7. Графики изменения эмпирических частот пропусков по "естественным" разрядам: а) в выборке студентов; б) в выборке артистов балета

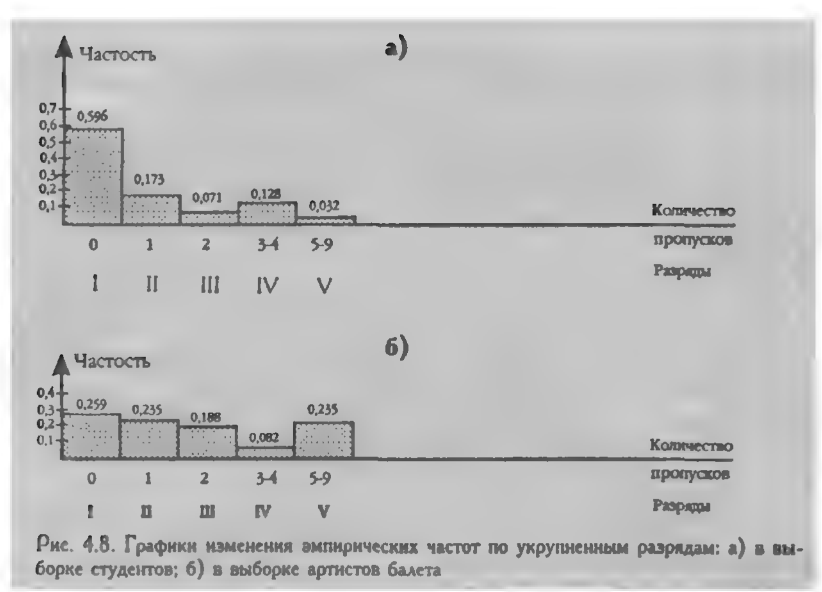


Рис. 4.8. Графики изменения эмпирических частот по укрупненным разрядам: а) в выборке студентов; б) в выборке артистов балета

Если бы у нас было больше испытуемых в выборке артистов балета, то, возможно, удалось бы сохранить подъем частоты на 5-и пропусках. Сейчас же нам придется довольствоваться сопоставлением по данным укрупненным разрядам.

Перейдем к подсчету теоретических частот для каждой ячейки Табл. 4.14

$f_{А\ теор}$=115*156/241=74,44

$f_{Б\ теор}$=115*85/241=40,56

$f_{В\ теор}$=47*156/241=30,41

$f_{Г\ теор}$=47*85/241=16,59

$f_{Д\ теор}$=27*156/241=17,47

$f_{Е\ теор}$=27*85/241=9,53 $f_{Ж\ теор}$=27*156/241=17,47

$f_{З\ теор}$=27*85/241=9,53 $f_{И\ теор}$=25*156/241=16,18 $f_{К\ теор}$=25*85/241=8,82

Определим количество степеней свободы V по формуле: $\nu=(k-1)*(c-1)$ где k - количество строк (разрядов),

с - количество столбцов (выборок). Для данного случая: ν=(5-1)*(2-1)=4

Все дальнейшие расчеты произведем в таблице по Алгоритму 13. Поправка на непрерывность не требуется, так как ν>1.

Таблица 4.15

Расчет критерия χ2 при сопоставлении двух эмпирических распределений пропусков слов в тесте Мюнстерберга (n_1=156, n_2=85)

Ячейки таблицы частот	Эмпирическая частота взгляда ($f_{эj}$)	Теоретическая частота ($f_т$)	($f_{эj}$-$f_т$)	$(f_{эj}-f_т)^2$	$(f_{эj}-f_т)^2/f_т$
А	93	74,44	18,56	344,47	4,63
Б	22	46,56	-18,56	344,47	8,49
В	27	30,41	-3,41	11,63	0,38
Г	20	16,59	3,41	11,63	0,70
Д	11	17,47	-6,47	41,86	2,40
Е	16	9,53	6,47	41,86	4,40
Ж	20	17,47	2,53	6,401	0,37
З	7	9,53	-2,53	6,401	0,67
И	5	16,18	-11,18	124,99	7,72
К	20	8,82	11,18	124,99	14,17
Суммы	241	241	0,00		43,95

По Табл. IX Приложения 1 определяем критические значения при ν =4:

$$\chi^2_{кр.} = \begin{cases} 9{,}488\ (\rho \le 0{,}05) \\ 13{,}277\ (\rho \le 0{,}01) \end{cases}$$

$\chi^2_{эмп}$=43,95

$\chi^2_{эмп} > \chi^2_{кр.}$

Ответ: H_0 отвергается. Принимается H_1. Распределения про-пусков слов в выборках студентов и артистов балета различаются между собой ($p<0,01$).

В распределении ошибок у артистов балета можно заметить два выраженных максимума (0 пропусков и 5 пропусков), что может указывать на два возможных источника ошибок[18].

4.3. λ - критерий Колмогорова-Смирнова

Назначение критерия

Критерий *X* предназначен для сопоставления двух распределений:

а) *эмпирического* с *теоретическим,* например, равномерным или нормальным;

б) одного *эмпирического* распределения с другими *эмпирическим* распределением.

Критерий позволяет найти точку, в которой сумма *накопленных расхождений* между двумя распределениями является наибольшей, и оценить достоверность этого расхождения.

Описание критерия

Если в методе $\chi 2$ мы сопоставляли частоты двух распределений отдельно по каждому разряду, то здесь мы сопоставляем сначала частоты по первому разряду, потом по сумме первого и второго разрядов, потом по сумме первого, второго и третьего разрядов и т. д. Таким образом, мы сопоставляем всякий раз накопленные к данному разряду частоты.

Если различия между двумя распределениями существенны, то в какой-то момент разность накопленных частот достигнет критического значения, и мы сможем признать различия статистически достоверными. В формулу критерия λ включается эта разность. Чем больше эмпирическое значение *λ,* тем более существенны различия.

Гипотезы

H_0: Различия между двумя распределениями недостоверны (судя по точке максимального накопленного расхождения между ними).

H_1: Различия между двумя распределениями достоверны (судя по точке максимального накопленного расхождения между ними).

Графическое представление критерия

Рассмотрим для иллюстрации распределение желтого (№4) цвета в 8-цветном тесте М. Люшера. Если бы испытуемые случайным образом выбирали цвета, то желтый цвет, так же, как и все остальные, равновероятно мог бы занимать любую из 8-и позиций выбора. На практике, однако, большинство испытуемых помещают этот цвет, "цвет ожидания и надежды" на одну из первых позиций ряда.

На Рис. 4.9 столбиками представлены относительные частоты[19] попадания желтого цвета сначала на 1-ю позицию (первый левый столбик), затем на 1-ю и 2-ю позицию (второй столбик), затем на 1-ю, 2-ю и 3-ю позиции и т. д. Мы видим, что высота столбиков постоянно возрастает, так как они отражают относительные частоты, накопленные к данной позиции. Например, столбик на 3-й позиции имеет высоту 0,51. Это означает, что

---

[18] Целесообразно было бы проверить совпадение распределения ошибок в обеих выборках с распределением Пуассона. Закону Пуассона подчиняются распределения редких событий, приходящихся 0, 1, *2,...* раз на сотни и тысячи наблюдений. Однако в данном случае эта модель неприменима: средняя и дисперсия не равны друг другу и составляют, соответственно, 0,91 и 1,96 в первой выборке и 2,29 и 5,43 во второй выборке.

[19] Относительная частота, или частость, - это частота, отнесенная к общему количеству наблюдений; в данном случае это частота попадания желтого цвета на дан-позицию, отнесенная к количеству испытуемых. Например, частота попадания желтого цвета на 1-ю позицию f=24; количество испытуемых n=102; относительная а $f^*=f/n$=0,235.

на первые три позиции желтый цвет помещают 51% испытуемых.

Прерывистой линией на Рис. 4.9 соединены точки, отражающие накопленные частоты, которые наблюдались бы, если бы желтый цвет с равной вероятностью попадал на каждую из 8-и позиций. Сплошными линиями обозначены расхождения между эмпирическими и теоретическими относительными частотами. Эти расхождения обозначаются как *d*.

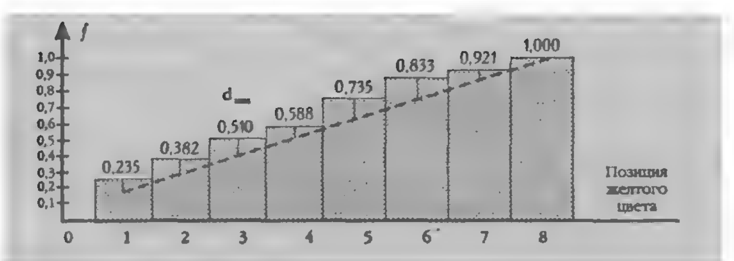


Рис. 4.9. Сопоставления в критерии *λ:* стрелками отмечены расхождения между эмпирическими и теоретическими накопленными относительными частотами по каждому разряду

Максимальное расхождение на Рис. 4.9 обозначено как *dmax.* Именно эта, третья позиция цвета, и является переломной точкой, определяющей, достоверно ли отличается данное эмпирическое распределение от равномерного. Мы проверим это при рассмотрении Примера 1.

Ограничения критерия *λ*

1. Критерий требует, чтобы выборка была достаточно большой. При сопоставлении двух эмпирических распределений необходимо, чтобы $n1,2 \geq 50$. Сопоставление эмпирического распределения с теоретическим иногда допускается при $n>5$ (Ван дер Варден Б.Л., 1960; Гублер Е.В., 1978).

2. Разряды должны быть упорядочены по нарастанию или убыванию какого-либо признака. Они обязательно должны отражать какое-то однонаправленное его изменение. Например, мы можем за разряды принять дни недели, 1-й, 2-й, 3-й месяцы после прохождения курса терапии, повышение температуры тела, усиление чувства недостаточности и т. д. В то же время, если мы возьмем разряды, которые случайно оказались выстроенными в данную последовательность, то и накопление частот будет отражать лишь этот элемент случайного соседства разрядов. Например, если шесть стимульных картин в методике Хекхаузена разным испытуемым предъявляются в разном порядке, мы не вправе говорить о накоплении реакций при переходе от картины №1 стандартного набора к картине №2 и т. д. Мы не можем говорить об однонаправленном изменении признака при сопоставлении категорий "очередность рождения", "национальность", "специфика полученного образования" и т.п. Эти данные представляют собой номинативные шкалы: в них нет никакого однозначного однонаправленного изменения признака.

Итак, мы не можем накапливать частоты по разрядам, которые отличаются лишь качественно и не представляют собой шкалы порядка. Во всех тех случаях, когда разряды представляют собой не упорядоченные по возрастанию или убыванию какого-либо признака категории, нам следует применять метод $\chi 2$.

Пример 1: Сопоставление эмпирического распределения с теоретическим

В выборке здоровых лиц мужского пола, студентов технических и военно-технических вузов в возрасте от 19-ти до 22 лет, средний возраст 20 лет, проводился тест Люшера в 8-цветном варианте. Установлено, что желтый цвет предпочитается испытуемыми чаще, чем отвергается (Табл. 4.16). Можно ли утверждать, что распределение желтого цвета по 8-и позициям у здоровых испытуемых отличается от равномерного распределения?

Таблица 4.16

Эмпирические частоты попадания желтого цвета на каждую из 8 позиций (n=102)

Разряды	Позиции желтого цвета								Сумма
	1	2	3	4	5	6	7	8	
Эмпирические частоты	24	25	13	8	15	10	9	8	102

Сформулируем гипотезы.

H_0: Эмпирическое распределение желтого цвета по восьми позициям не отличается от равномерного распределения.

H_1: Эмпирическое распределение желтого цвета по восьми позициям отличается от равномерного распределения.

Теперь приступим к расчетам, постепенно заполняя результатами таблицу расчета критерия λ. Все операции лучше прослеживать по Табл. 4.17, тогда они будут более понятными.

Занесем в таблицу наименования (номера) разрядов и соответствующие им эмпирические частоты (первый столбец Табл. 4.17). Затем рассчитаем эмпирические частости f^* по формуле:

$f^*_j = f_j / n$

где f_j - частота попадания желтого цвета на данную позицию;

n - общее количество наблюдений; j - номер позиции по порядку.

Запишем результаты во второй столбец (см. Табл. 4.17). Теперь нам нужно подсчитать накопленные эмпирические частости Σf^*. Для этого будем суммировать эмпирические частости f^*. Например, для 1-го разряда накопленная эмпирическая частость будет равняться эмпирической частости 1-го разряда, Σf^*_1=0,235[20].

Для 2-го разряда накопленная эмпирическая частость будет представлять собой сумму эмпирических частостей 1-го и 2-го разрядов:

Σf^*_{1+2}=0,235+0,147=0,382

Для 3-го разряда накопленная эмпирическая частость будет представлять собой сумму эмпирических частостей 1-го, 2-го и 3-го разрядов:

Σf^*_{1+2+3}=0,235+0,147+0,128=0,510

Мы видим, что можно упростить задачу, суммируя накопленную эмпирическую частость предыдущего разряда эмпирической частостью данного разряда, например, для 4-го разряда:

$\Sigma f^*_{1+2+3+4}$=0,510+0,078=0,588

Запишем результаты этой работы в третий столбец.

Теперь нам необходимо сопоставить накопленные эмпирические частости с накопленными теоретическими частостями. Для 1-го разряда теоретическая частость определяется по формуле:

$f^*_{теор}=1/k$

где k - количество разрядов (в данном случае - позиций цвета).

Для рассматриваемого примера:

$f^*_{теор}$=1/8=0,125

Эта теоретическая частость относится ко всем 8-и разрядам. Действительно, вероятность попадания желтого (или любого другого) цвета на каждую из 8-и позиций при случайном выборе составляет 1/8, т.е. 0,125.

Накопленные теоретические частости для каждого разряда определяем суммированием. Для 1-го разряда накопленная теоретическая частость равна теоретической частости попадания в разряд:

$f^*_{т1}$=0,125

---

[20] Все формулы приведены для дискретных признаков, которые могут быть выражены целыми числами, например: порядковый номер, количество испытуемых, количественный состав группы и т.п.

Для 2-го разряда накопленная теоретическая частость представляет собой сумму теоретических частостей 1-го и 2-го разрядов: $f^*_{т\ 1+2}$=0,125+0,125=0,250

Для 3-го разряда накопленная теоретическая частость представ ляет собой сумму накопленной к предыдущему разряду теоретической частости с теоретической частостью данного разряда:

$f^*_{т\ 1+2+3}$=0,250+0,125=0,375

Можно определить теоретические накопленные частости и путем! умножения:

$Sf^*_{т\ j}=f^*_{теор\ j}$

где $f^*_{теор}$ - теоретическая частость; j - порядковый номер разряда.

Занесем рассчитанные накопленные теоретические частости в четвертый столбец таблицы (Табл. 4.17).

Теперь нам осталось вычислить разности между эмпирическими и теоретическими накопленными частостями (столбцы 3-й и 4-й). В пятый столбец записываются абсолютные величины этих разностей, обозначаемые как *d*.

Определим по столбцу 5, какая из абсолютных величин разности является наибольшей. Она будет называться d_{max}. В данном случае d_{max}=0,135.

Теперь нам нужно обратиться к Табл. X Приложения 1 для определения критических значений d_{max} при *n*=102.

Таблица 4.17

Расчет критерия при сопоставлении распределения выборов желтого цвета с равномерным распределением (*n*=102)

Позиция желтого цвета	Эмпирическая частота	Эмпирическая частость	Накопленная эмпирическая частость	Накопленная теоретическая частость	Разность
1	24	0,235	0,235	0,125	0,110
2	15	0,147	0,382	0,250	0,132
3	13	0,128	0,510	0,375	**0,135**
4	8	0,078	0,588	0,500	0,088
5	15	0,147	0,735	0,625	0,110
6	10	0,098	0,833	0,750	0,083
7	9	0,088	0,921	0,875	0,046
8	8	0,079	1,000	1,000	0,000
Суммы	102	1,000			

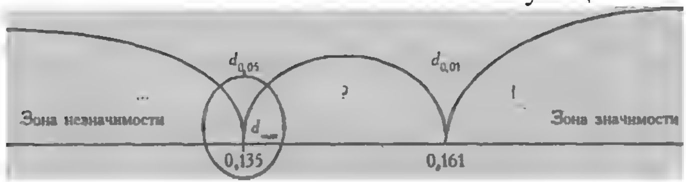
$$d_{кр}=\begin{cases} 1,36/\sqrt{n} & (\rho \le 0,05) \\ 1,63/\sqrt{n} & (\rho \le 0,01) \end{cases}$$

Для данного случая, следовательно,

$$d_{кр}=\begin{cases} 1,36/\sqrt{102} = 0,135 & (\rho \le 0,05) \\ 1,63/\sqrt{102} = 0,161 & (\rho \le 0,01) \end{cases}$$

Очевидно, что чем больше различаются распределения, тем больше и различия в накопленных частостях. Поэтому нам не составит труда распределить зоны значимости и незначимости по соответствующей оси:

Очевидно, что чем больше различаются распределения, тем больше и различия в накопленных частостях. Поэтому нам не составит труда распределить зоны значимости и незначимости по соответствующей оси:



$d_{эмп}$=0,135

$d_{эмп}$=dкр.

Ответ: H_0 отвергается при p=0,05. Распределение желтого цвета по восьми позициям отличается от равномерного распределения. Представим все выполненные

действия в виде алгоритма

АЛГОРИТМ 14

Расчет абсолютной величины разности *d* между эмпирическим и равномерным распределениями

1. Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты (первый столбец).

2. Подсчитать относительные эмпирические частоты (частости) для каждого разряда по формуле:

$f^*_{эмп}=f_{эмп}/n$

где $f_{эмп}$ - эмпирическая частота по данному разряду;

n - общее количество наблюдений. Занести результаты во второй столбец.

3. Подсчитать накопленные эмпирические частости Σf^*j по формуле:

где $\Sigma f^*j=\Sigma f^*_{j-1}+f^*_j$ - частость, накопленная на предыдущих разрядах; j - порядковый номер разряда; f^*_j- эмпирическая частость данного /-го разряда. Занести результаты в третий столбец таблицы.

4. Подсчитать накопленные теоретические частости для каждого раз ряда по формуле:

$\Sigma f^*_{т\ j}=\Sigma f^*_{т\ j-1}+f^*_{т\ j}$ где $\Sigma f^*_{т\ j-1}$ - теоретическая частость, накопленная на предыдущих

разрядах;

j - порядковый номер разряда;

$f^*_{т\ j}$ - теоретическая частость данного разряда. Занести результаты в третий столбец таблицы.

5. Вычислить разности между эмпирическими и теоретическими нако пленными частостями по каждому разряду (между значениями 3-го и 4-го столбцов).

6. Записать в пятый столбец абсолютные величины полученных раз ностей, без их знака. Обозначить их как *d*.

7. Определить по пятому столбцу наибольшую абсолютную величину разности - d_{max}.

8. По Табл. X Приложения 1 определить или рассчитать критические значения d_{max} для данного количества наблюдений *n*.

Если d_{max} равно критическому значению *d* или превышает его, различия между распределениями достоверны.

Пример 2: сопоставление двух эмпирических распределений

Интересно сопоставить данные, полученные в предыдущем примере, с данными обследования X. Кларом 800 испытуемых (Klar H., 1974, р. 67). X. Кларом было показано, что желтый цвет является единственным цветом, распределение которого по 8 позициям не отличается от равномерного. Для сопоставлений им использовался метод %

Полученные им эмпирические частоты представлены в Табл. 4.18.

Таблица 4.18

Эмпирические частоты попадания желтого цвета на каждую из 8 позиций в исследовании X. Клара (по: Klar H., 1974) (n=800)

Разряды-позиции желтого цвета	1	2	3	4	5	6	7	8	Сумма
Эмпирические частоты	98	113	116	87	91	112	97	86	800

Сформулируем гипотезы.

H_0: Эмпирические распределения желтого цвета по 8 позициям в отечественной выборке и выборке X. Клара не различаются.

H_1: Эмпирические распределения желтого цвета по 8 позициям в отечественной выборке и выборке X. Клара отличаются друг от друга.

Поскольку в данном случае мы будем сопоставлять накопленные эмпирические частости по каждому разряду, теоретические частости нас не интересуют.

Все расчеты будем проводить в таблице по алгоритму 15.

АЛГОРИТМ 15

Расчет критерия λ при сопоставлении двух эмпирических распределений

1. Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты, полученные в распределении 1 (первый столбец) и в распределении 2 (второй столбец).

2. Подсчитать эмпирические частости по каждому разряду для распределения 1 по формуле:

$f^*_э = f_э / n_1$

где $f_э$ - эмпирическая частота в данном разряде;

n_1 - количество наблюдений в выборке. Занести эмпирические частости распределения 1 в третий столбец.

3. Подсчитать эмпирические частости по каждому разряду для распределения 2 по формуле:

$f^*_э = f_э / n_2$

где $f_э$ - эмпирическая частота в данном разряде;

n_2 - количество наблюдений во 2-й выборке.

Занести эмпирические частости распределения 2 в четвертый столбец таблицы.

4. Подсчитать накопленные эмпирические частости для распределения 1 по формуле:

$$\sum f^*_i = \sum f^*_{i\ 1} + f^*_i$$

где Σf^*_{j-1} - частость, накопленная на предыдущих разрядах;

j - порядковый номер разряда;

f^*_{j-1}- частость данного разряда.

Полученные результаты записать в пятый столбец.

5. Подсчитать накопленные эмпирические частости для распределения 2 по той же формуле и записать результат в шестой столбец.

6. Подсчитать разности между накопленными частостями по каждому разряду. Записать в седьмой столбец абсолютные величины разностей, без их знака. Обозначить их как d.

7. Определить по седьмому столбцу наибольшую абсолютную величину разности

8. Подсчитать значение критерия λ по формуле:

$$\lambda = d_{max} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

где n_1 - количество наблюдений в первой выборке;

n_2 - количество наблюдений во второй выборке.

9. По Табл. XI Приложения 1 определить, какому уровню статистической значимости соответствует полученное значение λ.

Если $\lambda_{эмп} \geq 1{,}36$, различия между распределениями достоверны.

Последовательность выборок может быть выбрана произвольно, так как расхождения между ними оцениваются по абсолютной величине разностей. В нашем случае первой будем считать отечественную выборку, второй - выборку Клара.

Таблица 4.19

Расчет критерия при сопоставлении эмпирических распределений желтого цвета в отечественной выборке (n_1=102) и выборке Клара (n_2=800)

Позиция желтого цвета	Эмпирические частоты		Эмпирические частости		Накопленные эмпирические частости		Разность
	f_1	f_2	f^*_1	f^*_2	Σf^*_1	Σf^*_2	$\Sigma f^*_1 - \Sigma f^*_2$
1	24	98	0,235	0,123	0,235	0,123	0,112
2	15	113	0,147	0,141	0,382	0,264	0,118
3	13	116	0,128	0,145	0,510	0,409	0,101
4	8	87	0,078	0,109	0,588	0,518	0,070
5	15	91	0,147	0,114	0,735	0,632	0,103
6	10	112	0,098	0,140	0,833	0,772	0,061
7	9	97	0,088	0,121	0,921	0,893	0,028
8	8	86	0,079	0,107	1,000	1,000	0
Суммы	102	800	1,000	1,000			

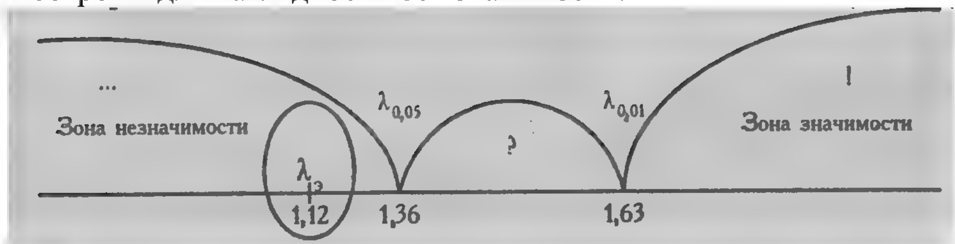
Максимальная разность между накопленными эмпирическими частостями составляет 0,118 и падает на второй разряд.

В соответствии с пунктом 8 алгоритма 15 подсчитаем значение Я,:

$$\lambda_{эмп} = 0{,}118 \cdot \sqrt{\frac{102 \cdot 800}{102 + 800}} = 1{,}12$$

По Табл. XI Приложения 1 определяем уровень статистической значимости полученного значения: р=0,16

Построим для наглядности ось значимости.



На оси указаны критические значения λ, соответствующие принятым уровням значимости: $\lambda_{0,05}$=1,*36*, $\lambda_{0,01}$=1,63.

Зона значимости простирается вправо, от 1,63 и далее, а зона незначимости - влево, от 1,36 к меньшим значениям.

$\lambda_{эмп} > \lambda_{кр}$

Ответ: H_0 принимается. Эмпирические распределения желтого цвета по 8 позициям в отечественной выборке и выборке X. Клара совпадают. Таким образом, распределения желтого цвета в двух выбор-ках не различаются, но в то же время они по-разному соотносятся с равномерным распределением: у Клара отличий от равномерного распределения не обнаружено, а в отечественной выборке различия обнаружены (ρ<0,05). Возможно, картину могло бы прояснить применение другого метода?

Е.В. Гублер (1978) предложил сочетать использование критерия λ, с критерием φ* (угловое преобразование Фишера).

Об этих возможностях сочетания методов λ и φ* мы поговорим в следующей главе (см. пример 4 п. 5.2).

4.4. Задачи для самостоятельной работы .

ВНИМАНИЕ!
При выборе способа решения задачи рекомендуется пользоваться
АЛГОРИТМОМ 16

Задача 6

В проективной методике X. Хекхаузена (модификация ТАТ) испытуемому последовательно предъявляются 6 картин. Всякий раз он сначала рассматривает картину в течение 20 сек, а затем в течение 5 минут пишет по ней рассказ, стараясь, в соответствии с инструкцией, проявить "максимум фантазии и воображения". После того, как испытуемый закончит писать первый рассказ, ему предъявляется вторая картина, и т. д. В данном исследовании разным испытуемым картины предъявлялись в разном порядке, так что каждая картина оказывалась первой, второй, третьей и т.д.

примерно одинаковое количество раз (Сидоренко Е. В., 1977).

При обследовании 113 студентов в возрасте от 20 до 35 лет; (средний возраст 23,2 года, 67 мужчин, 46 женщин) было установле-но, что в рассказах по картинам с условными названиями "Препо-даватель и ученик" и "Мастер измеряет деталь" словесные формулировки, отражающие "боязнь неудачи", встречаются гораздо чаще, чем в рассказах по другим картинам, в особенности по картине "Улыбающийся юноша" (см. Табл. 4.20).

Вопросы:

1) Можно ли утверждать, что картины методики обладают разной по будительной силой в отношении мотивов: а) "надежда на успех"; б) боязнь неудачи"?

2) Как следует из Табл. 4.20, нет почти ни одной картины, которая в равной мере стимулировала бы мотив "надежда на успех" и мотив "боязнь неудачи". Можно ли считать стимульный набор методики Хекхаузена неуравновешенным по направленности воздействия?

Таблица 4.20

Эмпирическое распределение словесных формулировок, отражающих мотивы "надежда на успех" и "боязнь неудачи" (*n*=113)

	Название картины		Количество вербальных реакций, отражающих "надежду на успех"	Количество вербальных реакций, отражающих "боязнь неудачи"		Всего
1	"Мастер измеряет деталь"	А	106	138	Б	244
2	"Преподаватель и ученик"	В	102	180	Г	282
3	"В цехе у машины"	А	108	34	Е	142
4	"У двери директора"	Ж	50	87	З	137
5	"Человек в бюро"	И	99	57	К	156
6	"Улыбающийся юноша"	Л	115	20	М	135
	Всего		580	516		1096

Задача 7

В процессе проведения транзактно-аналитических сессий установлено, что запреты на "психологические поглаживания[21]" встречаются с неодинаковой частотой. Например, многие участники тренинга признают у себя запрет "Не проси психологических поглаживаний у других людей", а запрет "Не давай психологических поглаживаний самому себе" встречается гораздо реже (см. Табл. 4.21).

Таблица 4.21

Частота встречаемости запретов на психологические поглаживания (n=166)

[21] Психологическое поглаживание - это "...любой акт, предполагающий признание присутствия другого человека" (Берн Э., 1992, с. 10). Практически в транзактно-аналитических сессиях под поглаживанием понимается выражение симпатии, восхищения, одобрения, любое искреннее признание положительных качеств и проявлений другого человека, к которым могут относиться внешние данные, глубинные личностные свойства, мастерство в своем деле, способность дарить психологическое тепло, и вовремя произнесенное слово и т.д.

	Запрет	Частота	Доля по отношению к общему количеству
1	Не давай психологических поглаживаний	44	15,66%
2	Не принимай психологических поглаживаний	45	16,01%
3	Не проси психологических поглаживаний	98	34,88%
4	Не отказывайся от психологических поглаживаний, даже если они тебе не нравятся	58	20,64%
5	Не давай психологических поглаживаний самому себе	36	12,81%
	Всего	281	100,00%

Вопросы:

Можно ли считать, что распределение запретов не является равномерным?

Можно ли утверждать, что запрет "Не проси" встречается достоверно чаще остальных?

Задача 8

В социально-психологическом исследовании стереотипов мужественности Н. В. Стан (1992) выборке из 31 женщин с высшим образованием в возрасте от 22 до 49 лет (средний возраст 35 лет) предъявлялись напечатанные на отдельных карточках перечни качеств, характеризующих один из четырех типов мужественности: мифологический, национальный, современный и религиозный. Испытуемым предлагалось внимательно ознакомиться с предложенными описаниями и выбрать из них то, которое в большей степени соответствует их представлению об идеальном мужчине. Затем испытуемым предлагалось выбрать одну из 3 оставшихся карточек, а затем одну из двух оставшихся. Результаты эксперимента представлены в Табл. 4.22.

Таблица 4.22

Распределение частот предпочтений 4 типов мужественности

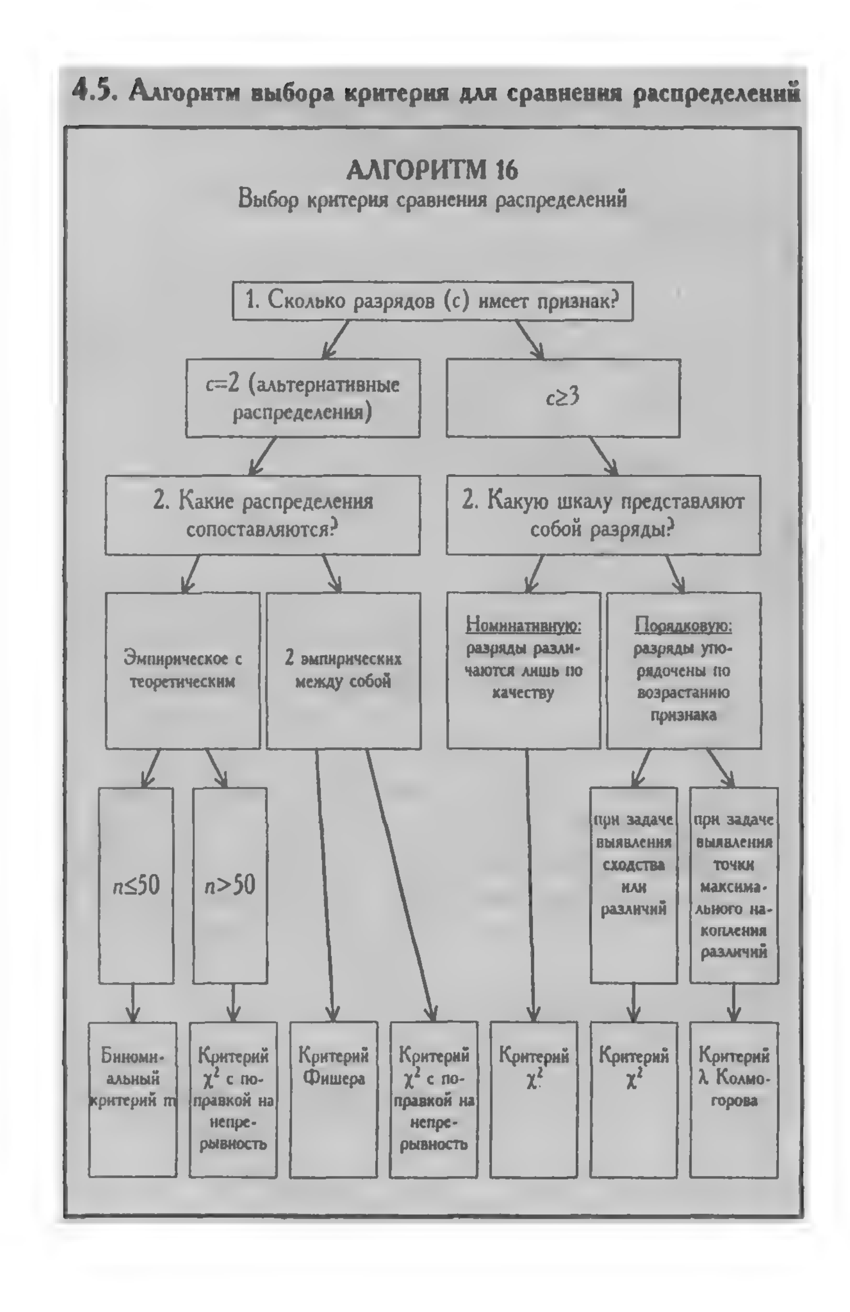
Тип мужественности	Эмпирические позиции				Всего
	1	2	3	4	
1. *Мифологический тип*: "Мощный, сильный, стройный, ловкий, бесстрашный, гордый, непокорный, уверенный, дерзкий, непреклонный, вспыльчивый, гневный, борец"	2 А	6 Б	4 В	19 Г	31
2. *Национальный тип*: "Ловкий, решительный, сдержанный, великодушный, преданный, открытый, бесхитростный, милосердный, уверенный, честный, доверчивый, защитник"	19 Д	4 Е	7 Ж	1 З	31
3. *Современный тип*: "Сильный, властный, сдержанный, уверенный, рассудочный, постоянный, агрессивный, практичный, эрудированный, самостоятельный, решительный, деятельный, энергичный, волевой"	7 И	10 К	12 Л	2 М	31
4. *Религиозный тип*: "Мягкий, миролюбивый, спокойный, кроткий, уступчивый, искренний, внимательный, выносливый, терпеливый, чувствительный"	3 Н	11 О	8 П	9 Р	31
Всего	31	31	31	31	124

Вопросы:

1) Различаются ли распределения предпочтений, выявленные по каждому из 4-х типов, между собой?

2) Можно ли утверждать, что предпочтение отдается какому-то одному или двум из типов мужественности? Наблюдается ли какая-либо групповая тенденция предпочтений?

4.5. Алгоритм выбора критерия для сравнения распределений



ГЛАВА 5
МНОГОФУНКЦИОНАЛЬНЫЕ СТАТИСТИЧЕСКИЕ КРИТЕРИИ

5.1. Понятие многофункциональных критериев

Многофункциональные статистические критерии - это критерии, которые могут использоваться по отношению к самым разнообразным | данным, выборкам и задачам.

Это означает, что данные могут быть представлены в любой шкале, начиная от номинативной (шкалы наименований).

Это означает также, что выборки могут быть как независимыми, так и "связанными", то есть мы можем с помощью многофункциональных критериев сравнивать и разные выборки испытуемых, и показатели одной и той же выборки, измеренные в разных условиях. Нижние границы выборок - 5 наблюдений, но возможно применение критериев и по отношению к выборкам с *n=2,* с некоторыми оговорками (см. разделы "Ограничения критерия φ*" и "Ограничения биномиального критерия m”)

Верхняя граница выборок задана только в биномиальном критерии - 50 человек. В критерии φ* Фишера верхней границы не существует - выборки могут быть сколь угодно большими.

Многофункциональные критерии позволяют решать задачи сопоставления уровней исследуемого признака, сдвигов в значениях исследуемого признака и сравнения распределений.

К числу многофункциональных критериев в полной мере относится критерий φ* Фишера (угловое преобразование Фишера) и, с некоторыми оговорками - биномиальный критерий m.

Многофункциональные критерии построены на сопоставлении долей, выраженных в долях единицы или в процентах. Суть критериев [состоит в определении того, какая доля наблюдений (реакций, выборов, испытуемых) в данной выборке характеризуется интересующим исследователя эффектом и какая доля этим эффектом не характеризуется.

Таким эффектом может быть:

а) определенное *значение* качественно определяемого признака - например, выражение согласия с каким-либо предложением; выбор правой дорожки из двух симметричных дорожек; отнесенность к определенному полу; присутствие фигуры отца в раннем воспоминании и др.

б) определенный *уровень* количественно измеряемого признака, например, получение оценки, превосходящей проходной балл; решение задачи менее чем за 20 сек; факт работы в команде, по численности превышающей 4-х человек; выбор дистанции в разговоре, превышающей 50 см, и др.

в) определенное *соотношение* значений или уровней исследуемого признака, например, более частый выбор альтернатив А и Б по сравнению с альтернативами В и Г; преимущественное проявление крайних значений признака, как самых высоких, так и самых низких; преобладание положительных сдвигов над отрицательными и др.

Итак, путем сведения любых данных к альтернативной шкале "Есть эффект - нет аффекта" многофункциональные критерии позволяют решать все три задачи сопоставлений - сравнения "уровней", оценки "сдвигов" и сравнения распределений.

Критерий φ* применяется в тех случаях, когда обследованы *две* выборки испытуемых, биномиальный критерий m - в тех случаях, когда обследована лишь *одна* выборка испытуемых. Правила выбора одного из этих критериев отражены в Алгоритме **19.**

5.2. Критерий φ* — угловое преобразование Фишера

Данный метод описан во многих руководствах (Плохинский Н.А., 1970; Гублер Е.В., 1978; Ивантер Э.В., Коросов А.В., 1992 и др.) Настоящее описание опирается на тот вариант метода, который был разработан и изложен Е.В. Гублером.

Назначение критерия φ*

Критерий Фишера предназначен для сопоставления двух выборок по частоте встречаемости интересующего исследователя эффекта.

Описание критерия

Критерий оценивает достоверность различий между процентными долями двух выборок, в которых зарегистрирован интересующий нас эффект.

Суть углового преобразования Фишера состоит в переводе процентных долей в величины центрального угла , который измеряется в радианах . Большей процентной доле будет соответствовать больший угол ф, а меньшей доле - меньший угол, но соотношения здесь не линейные:

$$\varphi = 2 \cdot \arcsin(\sqrt{P})$$

где P - процентная доля, выраженная в долях единицы (см. Рис. 5.1).

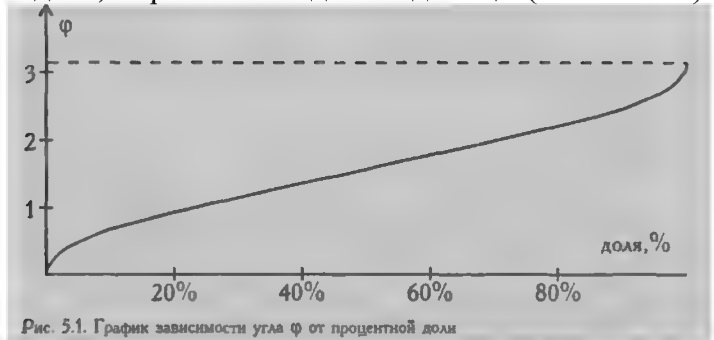


Рис. 5.1. График зависимости угла φ от процентной доли

При увеличении расхождения между углами φ_1 и φ_2 и увеличения численности выборок значение критерия возрастает. Чем больше величина φ* , тем более вероятно, что различия достоверны.

Гипотезы

H_0: Доля лиц, у которых проявляется исследуемый эффект, в выборке 1 не больше, чем в выборке 2.

H_1: Доля лиц, у которых проявляется исследуемый эффект, в выборке 1 больше, чем в выборке 2.

Графическое представление критерия φ*

Метод углового преобразования несколько более абстрактен, чем остальные критерии.

Формула, которой придерживается Е. В. Гублер при подсчете значений φ, предполагает, что 100% составляют угол φ=3,142, то есть округленную величину π=3,14159... Это позволяет нам представить сопоставляемые выборки в виде двух полукругов, каждый из которых символизирует 100% численности своей выборки. Процентные доли испытуемых с "эффектом" будут представлены как секторы, образованные центральными углами φ. На Рис. 5.2 представлены два полукруга, иллюстрирующие Пример 1. В первой выборке 60% испытуемых решили задачу. Этой процентной доле соответствует угол φ=1,772. Во второй выборке 40% испытуемых решили задачу. Этой процентной доле соответствует угол φ =1,369.

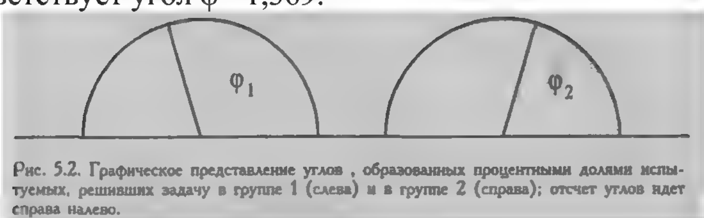


Рис. 5.2. Графическое представление углов , образованных процентными долями испытуемых, решивших задачу в группе 1 (слева) и в группе 2 (справа); отсчет углов идет справа налево.

Критерий φ* позволяет определить, действительно ли один из углов статистически достоверно превосходит другой при данных объемах выборок.

Ограничения критерия φ*

1. Ни одна из сопоставляемых долей не должна быть равной нулю. Формально нет препятствий для применения метода φ в случаях, когда доля наблюдений в одной из выборок равна 0. Однако в этих случаях результат может оказаться неоправданно завышенным (Гублер Е.В., 1978, с. 86).

2. Верхний предел в критерии φ отсутствует - выборки могут быть сколь угодно большими.

Нижний предел - 2 наблюдения в одной из выборок. Однако должны соблюдаться следующие соотношения в численности двух выборок:

а) если в одной выборке всего 2 наблюдения, то во второй должно быть не менее 30:

$n_1=2 \rightarrow n_2 \geq 30;$

б) если в одной из выборок всего 3 наблюдения, то во второй должно быть не менее 7:

$n_1=3 \rightarrow n_2 \geq 7;$

в) если в одной из выборок всего 4 наблюдения, то во второй должно быть не менее 5:

$n_1=4 \rightarrow n_2 \geq 5;$

г) при $n_1, n_2 \geq 5$ возможны любые сопоставления.

В принципе возможно и сопоставление выборок, не отвечающих этому условию, например, с соотношением n_1=2, n_2=15, но в этих случаях не удастся выявить достоверных различий.

Других ограничений у критерия φ* нет.

Рассмотрим несколько примеров, иллюстрирующих возможности критерия φ*.

Пример 1: сопоставление выборок по качественно определяемому признаку.

Пример 2: сопоставление выборок по количественно измеряемому признаку.

Пример 3: сопоставление выборок и по уровню, и по распределению признака.

Пример 4: использование критерия φ* в сочетании с критерием *X* Колмогорова-Смирнова в целях достижения максимально точного результата.

Пример 1 - сопоставление выборок по качественно определяемому признаку

В данном варианте использования критерия мы сравниваем процент испытуемых в одной выборке, характеризующихся каким-либо качеством, с процентом испытуемых в другой выборке, характеризующихся тем же качеством.

Допустим, нас интересует, различаются ли две группы студентов по успешности решения новой экспериментальной задачи. В первой группе из 20 человек с нею справились 12 человек, а во второй выборке из 25 человек - 10. В первом случае процентная доля решивших задачу составит 12/20·100%=60%, а во второй 10/25·100%=40%. Достоверно ли различаются эти процентные доли при данных n_1 и n_2?

Казалось бы, и "на глаз" можно определить, что 60% значительно выше 40%. Однако на самом деле эти различия при данных n_1, n_2 недостоверны.

Проверим это. Поскольку нас интересует факт решения задачи, будем считать "эффектом" успех в решении экспериментальной задачи, а отсутствием эффекта - неудачу в ее решении.

Сформулируем гипотезы.

H_0: Доля лиц, справившихся с задачей, в первой группе не больше, чем во второй группе.

H_1: Доля лиц, справившихся с задачей, в первой группе больше, чем во второй группе.

Теперь построим так называемую четырехклеточную, или четырехпольную таблицу, которая фактически представляет собой таблицу эмпирических частот по двум значениям признака: "есть эффект" - "нет эффекта".

Таблица 5.1

Четырехклеточная таблица для расчета критерия при сопоставлении двух групп испытуемых по процентной доле решивших задачу.

Группы	"Есть эффект": задача решена			"Нет эффекта": задача не решена			Суммы
	Количество испытуемых	% доля		Количество испытуемых	% доля		
1 группа	12	(60%)	А	8	(40%)	Б	20
2jJynna	10	(40%)	В	15	(60%)	Г	25
Суммы	22			23			45

В четырехклеточной таблице, как правило, сверху размечаются столбцы "Есть эффект" и "Нет эффекта", а слева - строки "1 группа" и "2 группа". Участвуют в сопоставлениях, собственно, только поля (ячейки) А и В, то есть процентные доли по столбцу "Есть эффект".

По Табл. XII Приложения 1 определяем величины φ, соответствующие процентным долям в каждой из групп.

$\varphi_{1(60\%)}=1{,}772$

$\varphi_{2(40\%)}=1{,}369$

Теперь подсчитаем эмпирическое значение φ* по формуле:

$$\varphi^* = (\varphi_1 - \varphi_2)\cdot\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

где φ_1 - угол, соответствующий большей % доле;

φ_2 - угол, соответствующий меньшей % доле;

n_1 - количество наблюдений в выборке 1;

n_2 - количество наблюдений в выборке 2.

В данном случае:

$$\varphi^*_{эмп} = (1{,}772 - 1{,}369)\cdot\sqrt{\frac{20\cdot 25}{20+25}} = 0{,}403\cdot\sqrt{11{,}11} = 1{,}34$$

По Табл. XIII Приложения 1 определяем, какому уровню значимости соответствует $\varphi^*_{эмп}=1{,}34$:

p=0,09

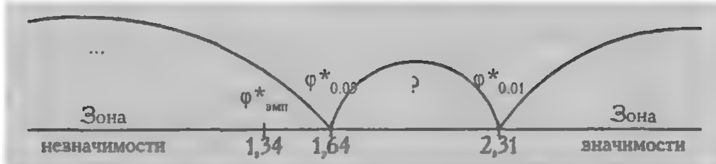
Можно установить и критические значения φ*, соответствующие принятым в психологии уровням статистической значимости:

$$\varphi^*_{кр} = \begin{cases} 1{,}64\ (p \le 0{,}05) \\ 2{,}31\ (p \le 0{,}01) \end{cases}$$

$\varphi_{эмп}=1{,}34$

$\varphi_{эмп}<\varphi_{кр}$

Построим "ось значимости".



Полученное эмпирическое значение φ* находится в зоне незначимости.

Ответ: H_0 принимается. Доля лиц, справившихся с задачей, в первой группе не больше, чем во второй группе.

Можно лишь посочувствовать исследователю, который считает существенными различия в 20% и даже в 10%, не проверив их достоверность с помощью критерия φ*. В данном случае, например, достоверными были бы только различия не менее чем в 24,3%.

Похоже, что при сопоставлении двух выборок по какому-либо качественному признаку критерий φ может нас скорее огорчить, чем обрадовать. То, что казалось

существенным, со статистической точки зрения может таковым не оказаться.

Гораздо больше возможностей порадовать исследователя появляется у критерия Фишера тогда, когда мы сопоставляем две выборки по количественно измеренным признакам и можем варьировать "эффект .

Пример 2 - сопоставление двух выборок по количественно измеряемому признаку

В данном варианте использования критерия мы сравниваем процент испытуемых в одной выборке, которые достигают определенного уровня значения признака, с процентом испытуемых, достигающих этого уровня в другой выборке.

В исследовании Г. А. Тлегеновой (1990) из 70 юношей - учащихся ПТУ в возрасте от 14 до 16 лет было отобрано по результатам обследования по Фрайбургскому личностному опроснику 10 испытуемых с высоким показателем по шкале Агрессивности и 11 испытуемых с низким показателем по шкале Агрессивности. Необходимо определить, различаются ли группы агрессивных и неагрессивных юношей по показателю расстояния, которое они спонтанно выбирают в разговоре с сокурсником. Данные Г. А. Тлегеновой представлены в Табл. 5.2. Можно заметить, что агрессивные юноши чаще выбирают расстояние в 50 см или даже меньше, в то время как неагрессивные юноши чаще выбирают расстояние, превышающее 50 см.

Теперь мы можем рассматривать расстояние в 50 см как критическое и считать, что если выбранное испытуемым расстояние меньше или равно 50 см, то "эффект есть", а если выбранное расстояние больше 50 см, то "эффекта нет". Мы видим, что в группе агрессивных юношей эффект наблюдается в 7 из 10, т. е. в 70% случаев, а в группе неагрессивных юношей - в 2 из 11, т. е. в 18,2% случаев. Эти процентные доли можно сопоставить по методу φ* , чтобы установить достоверность различий между ними.

Таблица 5.2

Показатели расстояния (в см), выбираемого агрессивными и неагрессивными юношами в разговоре с сокурсником (по данным Г.А. Тлегеновой, 1990)

	Группа 1: юноши с высокими показателями по шкале Агрессивности FPI-R[22] (n_1 =10)		Группа 2: юноши с низкими значениями по шкале Агрессивности FPI-R (n_2 =11)	
	d(см)	% доля	*d(см)*	% доля
"Есть эффект" d≤50 см	30 40 50 50 50 50 50	70%	40 45	18,2%
"Нет эффекта" d>50 см	70 80 QO	30%	65 75 75 75 75 100 100 100 100	. 81,8%
Суммы	560	100%	850	100%
Средние	5б:о		77.3	

[22] FPI-R - Фрайбургский личностный опросник

Сформулируем гипотезы.

H_0: Доля лиц, которые выбирают дистанцию *d≤50* см, в группе агрессивных юношей не больше, чем в группе неагрессивных юношей.

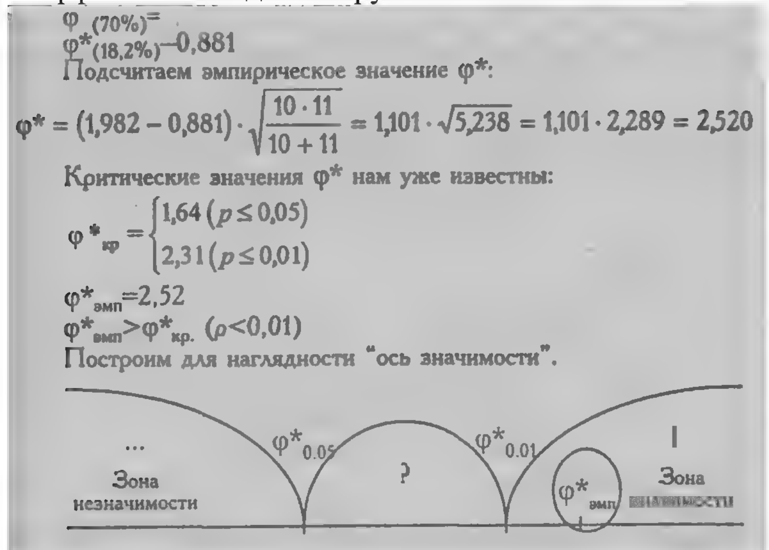
H_1: Доля лиц, которые выбирают дистанцию d≤50 см, в группе агрессивных юношей больше, чем в группе неагрессивных юношей. Теперь построим так называемую четырехклеточную таблицу.

Таблица 53

Четырехклеточная таблица для расчета критерия φ* при сопоставлении групп агрессивных (nf=10) и неагрессивных юношей (п2=11)

Группы	"Есть эффект": d≤50			"Нет эффекта". d>50			Суммы
	Количество испытуемых	(% доля)		Количество испытуемых	(% доля)		
1 группа - агрессивные юноши	7	(70%)	А	3	(30%)	Б	10
2 группа - неагрессивные юноши	2	(180%)	В	9	(81,8%)	Г	и
Сумма	9			12			21

По Табл. XII Приложения 1 определяем величины φ, соответствующие процентным долям "эффекта" в каждой из групп.


$\varphi_{(70\%)}=$

$\varphi^*_{(18,2\%)}$—0,881

Подсчитаем эмпирическое значение φ*:

$$\varphi^* = (1,982 - 0,881)\cdot\sqrt{\frac{10\cdot 11}{10+11}} = 1,101\cdot\sqrt{5,238} = 1,101\cdot 2,289 = 2,520$$

Критические значения φ* нам уже известны:

$$\varphi^*_{кр} = \begin{cases} 1,64\ (p\le 0,05) \\ 2,31\ (p\le 0,01) \end{cases}$$

$\varphi^*_{эмп}=2,52$

$\varphi^*_{эмп}>\varphi^*_{кр.}$ $(\rho<0,01)$

Построим для наглядности "ось значимости".

Полученное эмпирическое значение φ* находится в зоне значимости.

Ответ: H_0 отвергается. Принимается H_1. Доля лиц, которые выбирают дистанцию в беседе меньшую или равную 50 см, в .группе агрессивных юношей больше, чем в группе неагрессивных юношей

На основании полученного результата мы можем сделать заключение, что более агрессивные юноши чаще выбирают расстояние менее полуметра, в то время как неагрессивные юноши чаще выбирают большее, чем полметра, расстояние. Мы видим, что агрессивные юноши общаются фактически на границе интимной (0—46 см) и личной зоны (от 46 см). Мы помним, однако, что интимное расстояние между партнерами является прерогативой не только близких добрых отношений, но и рукопашного боя (Hall E.T., 1959).

Пример 3 - сопоставление выборок и по уровню, и по распределению признака.

В данном варианте использования критерия мы вначале можем проверить, различаются ли группы по уровню какого-либо признака, а затем сравнить распределения признака в двух выборках. Такая задача может быть актуальной при анализе различий в диапазонах или форме распределения оценок, получаемых испытуемыми по какой-либо

новой методике.

В исследовании Р. Т. Чиркиной (1995) впервые использовался опросник, направленный на выявление тенденции к вытеснению из памяти фактов, имен, намерений и способов действия, обусловленному личными, семейными и профессиональными комплексами. Опросник был создан при участии Е. В. Сидоренко на основании материалов книги 3. Фрейда "Психопатология обыденной жизни". Выборка из 50 студентов Педагогического института, не состоящих в браке, не имеющих детей, в возрасте от 17 до 20 лет, была обследована с помощью данного опросника, а также методики Менестера-Корзини для выявления интенсивности ощущения собственной недостаточности, или "комплекса неполноценности" (Manaster G. J., Corsini R. J., 1982).

Результаты обследования представлены в Табл. 5.4.

Можно ли утверждать, что между показателем энергии вытеснения, диагностируемым с помощью опросника, и показателями интенсивности, ощущения собственной недостаточности существуют какие-либо значимые соотношения?

Таблица 5.4

Показатели интенсивности ощущения собственной недостаточности в группах студентов с высокой (nj=18) и низкой (п2=24) энергией вытеснения

	Группа 1: энергия вытеснения от 19 до 31 балла (n_1=181	Группа 2: энергия вытеснения от 7 до 13 баллов (n_2=24)
	0; 0; 0; 0; 0	0; 0
		5; 5; 5; 5
		10; 10; 10; 10; 10; 10
		15; 15
	20; 20	20; 20; 20; 20
	30; 30; 30; 30; 30; 30; 30	30; 30; 30; 30; 30; 30
	50; 50	
	60; 60	
Суммы	470	370
Средние	26,11	15,42

Несмотря на то, что средняя величина в группе с более энергичным вытеснением выше, в ней наблюдаются также и 5 нулевых значений. Если сравнить гистограммы распределения оценок в двух выборках, то между ними обнаруживается разительный контраст (Рис. 5.3).

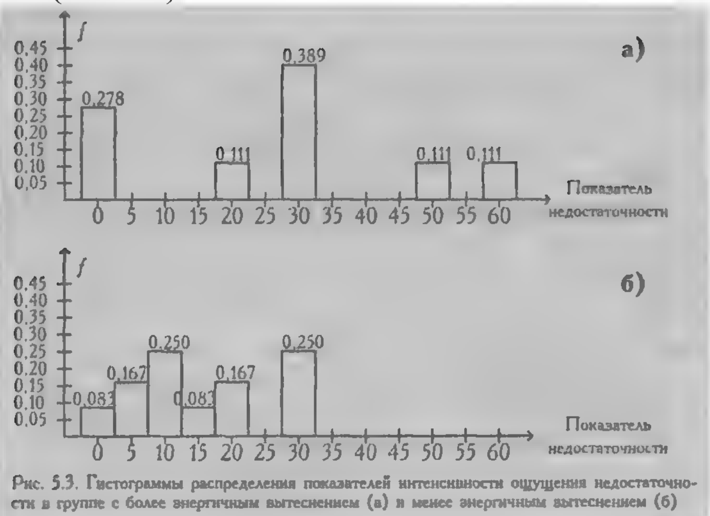


Рис. 5.3. Гистограммы распределения показателей интенсивности ощущения недостаточности в группе с более энергичным вытеснением (а) и менее энергичным вытеснением (б)

Для сравнения двух распределений мы могли бы применить критерий χ^2 или критерий λ, но для этого нам пришлось бы укрупнять разряды, а кроме того, в обеих выборках n<30.

Критерий φ* позволит нам проверить наблюдаемый на графике эффект несовпадения двух распределений, если мы условимся считать, что "эффект есть", если показатель чувства недостаточности принимает либо очень низкие (0), либо, наоборот, очень высокие значения (S30), и что "эффекта нет", если показатель чувства

недостаточности принимает средние значения, от 5 до 25.

Сформулируем гипотезы.

H_0: Крайние значения показателя недостаточности (либо 0, либо 30 и более) в группе с более энергичным вытеснением встречаются не чаще, чем в группе с менее энергичным вытеснением.

H_1: Крайние значения показателя недостаточности (либо 0, либо 30 и более) в группе с более энергичным вытеснением встречаются чаще, чем в группе с менее энергичным вытеснением.

Создадим четырехклеточную таблицу, удобную для дальнейшего расчета критерия φ*.

Таблица 5.5

Четырехклеточная таблица для расчета критерия φ*при сопоставлении групп с большей и меньшей энергией вытеснения по соотношению показателей недостаточности

Группы	"Есть эффект": показатель недостаточности равен 0 или >30	"Нет эффекта": показатель недостаточности от 5 до 25	Суммы
1 группа - с большей энергией вытеснения	16 (88,9%)	2 (11,1%)	18
2 группа - с меньшей энергией вытеснения	8 (33,3%)	16 (66,7%)	24
Суммы	24	18	42[23]

По Табл. XII Приложения 1 определим величины ф, соответствующие сопоставляемым процентным долям:

$\varphi_{1(88,9\%)} = 2{,}462$

$\varphi_{1(33,3\%)} = 1{,}230$

Подсчитаем эмпирическое значение φ*:

$$\varphi^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (2{,}462 - 1{,}23) \cdot \sqrt{\frac{18 \cdot 24}{18 + 24}} =$$
$$= 1{,}232 \cdot \sqrt{10{,}286} = 1{,}232 \cdot 3{,}207 = 3{,}951$$

Критические значения φ* при любых n_1, n_2, как мы помним из предыдущего примера, составляют:

$$\varphi^*_{кр.} = \begin{cases} 1{,}64 \ (p \le 0{,}05) \\ 2{,}31 \ (p \le 0{,}01) \end{cases}$$

$\varphi^*_{эмп} = 3{,}951$

$\varphi^*_{эмп} > \varphi^*_{кр.}$ $(\rho \le 0{,}01)$

Табл. XIII Приложения 1 позволяет нам и более точно определить уровень значимости полученного результата: p<0,001.

Ответ: H_0 отвергается. Принимается H_1. Крайние значения показателя недостаточности (либо 0, либо 30 и более) в группе с большей энергией вытеснения встречаются чаще, чем в группе с меньшей энергией вытеснения.

Итак, испытуемые с большей- энергией вытеснения могут иметь как очень высокие (30 и более), так и очень низкие (нулевые) показатели ощущения собственной недостаточности. Можно предположить, что они вытесняют и свою неудовлетворенность, и потребность в жизненном успехе. Эти предположения нуждаются в дальнейшей проверке.

Полученный результат, независимо от его интерпретации, подтверждает

[23] В первоначальной выборке было 50 человек, но 8 из них были исключены из рассмотрения как имеющие средний балл по показателю анергии вытеснения (14-15). Показатели интенсивности чувства недостаточности у них тоже средние: 6 значений по 20 баллов и 2 значения по 25 баллов.

возможности критерия φ* в оценке различий в форме распределения признака в двух выборках.

В мощных возможностях критерия φ* можно убедиться, подтвердив совершенно иную гипотезу при анализе материалов данного примера. Мы можем доказать, например, что в группе с большей энергией вытеснения показатель недостаточности все же выше, несмотря на парадоксальность его распределения в этой группе.

Сформулируем новые гипотезы.

H_0 Наиболее высокие значения показателя недостаточности (30 и более) в группе с большей энергией вытеснения встречаются не чаще, чем в группе с меньшей энергией вытеснения.

H_1*:* Наиболее высокие значения показателя недостаточности (30 и более) в группе с большей энергией вытеснения встречаются чаще, чем в группе с меньшей энергией вытеснения. Построим четырехпольную таблицу, используя данные Табл. 5.4.

Таблица 5.6

Четырехклеточная таблица для расчета критерия φ* при сопоставлении групп с большей и меньшей энергией вытеснения по уровню показателя недостаточности

Группы	"Есть эффект"* показатель недостаточности больше или равен 30		"Нет эффекта": показатель недостаточности меньше 30		Суммы
1 группа - с большей энергией вытеснения	11	(61,1%)	7	(38.9%)	18
2 группа - с меньшей энергией вытеснения	6	(25.0%)	18	(75.0%)	24
Суммы	17		25		42

По Табл. XII Приложения 1 определяем величины φ:

$\varphi_{1(61,1\%)}=1,795$

$\varphi_{2(25,0\%)}=1,047$

Подсчитываем эмпирическое значение φ*:

$$\varphi^* = (1,795 - 1,047) \cdot \sqrt{\frac{18 \cdot 24}{18 + 24}} = 0,748 \cdot 3,207 = 2,398$$

По Табл. XIII Приложения 1 определяем, что этот результат соответствует уровню значимости р=0,008.

Ответ: Но отвергается. Принимается Hj: Наиболее высокие показатели недостаточности (30 и более баллов) в группе с большей энергией вытеснения встречаются чаще, чем в группе с меньшей энергией вытеснения (р=0,008).

Итак, нам удалось доказать и то, что в группе с более энергичным вытеснением преобладают крайние значения показателя недостаточности, и то, что больших своих значений этот показатель достигает именно в этой группе.

Теперь мы могли бы попробовать доказать, что в группе с большей энергией вытеснения чаще встречаются и более низкие значения показателя недостаточности, несмотря на то, что средняя величина **в** этой группе больше (26,11 против 15,42 в группе **с** меньшим вытеснением).

Сформулируем гипотезы.

H_0: Самые низкие показатели недостаточности (нулевые) в группе **с** большей энергией вытеснения встречаются не чаще, чем в группе **с** меньшей энергией вытеснения.

H_1: Самые низкие показатели недостаточности (нулевые) встречаются **в** группе с большей энергией вытеснения чаще, чем в группе **с** менее энергичным вытеснением. Сгруппируем данные в новую четырехклеточную таблицу.

Таблица 5.7

Четырехклеточная таблица для сопоставления групп с разной энергией вытеснения по частоте нулевых значений показателя недостаточности

Группы	"Есть эффект": показатель недостаточности	"Нет эффекта" недостато	показатель не равен 0	Суммы

	равен 0	чности	
1 группа - с большей энергией вытеснения	5 (27,8%)	13 (72,2%)	18
1 группа - с меньшей энергией вытеснения	2 (8,3%)	22 (91,7%)	24
Суммы	7	35	42

Определяем величины φ и подсчитываем значение φ*:

$\varphi_{1(27,8\%)}=1,111$

$\varphi_{2(8,3\%)}=0,584$

$$\varphi^* = (1,111 - 0,584)\cdot\sqrt{\frac{18\cdot 24}{18+24}} = 0,527\cdot 3,207 = 1,69$$

$$\varphi^*_{кр} = \begin{cases} 1,64\,(p\le 0,05) \\ 2,31\,(p\le 0,01) \end{cases}$$

$\varphi^*_{эмп} > \varphi^*_{кр}$ $(p\le 0,05)$

Ответ: H_0 отвергается. Самые низкие показатели недостаточности (нулевые) в группе с большей энергией вытеснения встречаются чаще, чем в группе с меньшей энергией вытеснения (р<0,05).

В сумме полученные результаты могут рассматриваться как свидетельство частичного совпадения понятий комплекса у З.Фрейда и А.Адлера.

Существенно при этом, что между показателем энергии вытеснения и показателем интенсивности ощущения собственной недостаточности в целом по выборке получена положительная линейная корреляционная связь (р=+0,491, р<0,01). Как мы можем убедиться, применение критерия φ* позволяет проникнуть в более тонкие и содержательно значимые соотношения между этими двумя показателями.

Пример 4 - использование критерия φ* в сочетании с критерием λ Колмогорова-Смирнова в целях достижения максимально точного **результата**

Если выборки сопоставляются по каким-либо количественно измеренным показателям, встает проблема выявления той точки распределения, которая может использоваться как критическая при разделении всех испытуемых на тех, у кого "есть эффект" и тех, у кого "нет эффекта".

В принципе точку, по которой мы разделили бы группу на подгруппы, где есть эффект и нет эффекта, можно выбрать достаточно произвольно. Нас может интересовать любой эффект и, следовательно, мы можем разделить обе выборки на две части в любой точке, лишь бы это имело какой-то смысл.

Для того, чтобы максимально повысить мощность критерия φ*, нужно, однако, выбрать точку, в которой различия между двумя сопоставляемыми группами являются наибольшими. Точнее всего мы сможем сделать это с помощью алгоритма расчета критерия λ , позволяющего обнаружить точку максимального расхождения между двумя выборками.

Возможность сочетания критериев φ* и λ описана Е.В. Гублером (1978, с. 85-88). Попробуем использовать этот способ в решении следующей задачи.

В совместном исследовании М.А. Курочкина, Е.В. Сидоренко и Ю.А. Чуракова (1992) в Великобритании проводился опрос английских общепрактикующих врачей двух категорий: а) врачи, поддержавшие медицинскую реформу и уже превратившие свои приемные в фондодержащие организации с собственным бюджетом; б) врачи, чьи приемные по-прежнему не имеют собственных фондов и целиком обеспечиваются государственным бюджетом. Опросники были разосланы выборке из 200 врачей, репрезентативной по отношению к генеральной совокупности английских врачей по представленности лиц разного пола, возраста, стажа и места работы - в крупных городах

или в провинции.

Ответы на опросник прислали 78 врачей, из них 50 работающих в приемных с фондами и 28 - из приемных без фондов. Каждый из врачей должен был прогнозировать, какова будет доля приемных с фондами в следующем, 1993 году. На данный вопрос ответили только 70 врачей из 78, приславших ответы. Распределение их прогнозов представлено в Табл. 5.8 отдельно для группы врачей с фондами и группы врачей без фондов.

Различаются ли каким-то образом прогнозы врачей с фондами и врачей без фондов?

Таблица 5.8

Распределение прогнозов сбщепрактикующих врачей о том, какова будет доля приемных с фондами в 1993 году

Прогнозируемая доля приемных с фондами	Эмпирические частоты выбора данной категории прогноза		
	врачами с фондом (n_1=45)	врачами без фонда (n_2 =25)	Суммы
1. от 0 до 20%	4	5	9
2. от 21 до 40%	15	И	26
3. от 41 до 60%	18	5	23
4. от 61 до 80%	7	4	И
5. от 81 до 100%	1	0	1
Суммы	45	25	70

Определим точку максимального расхождения между двумя распределениями ответов по Алгоритму 15 из п. 4.3 (см. Табл. 5.9).

Таблица 5.9

Расчет максимальной разности накопленных частостей в распределениях прогнозов врачей двух групп

Прогнозируемая доля приемных с фондами (%)	Эмпирические частоты выбора данной категории ответа		Эмпирические частости		Накопленные эмпирические частости		Разность (d)
	врачами с фондом (n_1=45)	врачами без фонда (n_2=25)	$f^*_{э1}$	f^*_{a2}	$\sum f^*_{э1}$	$\sum f^*_{a1}$	
1. от 0 до 20%	4	5	0,089	0,200	0,089	0,200	0111
2. от 21 до 40%	15	11	0,333	0,440	0,422	0,640	**0,218**
3. от 41 до 60%	18	5	0,400	0,200	0,822	0,840	0,018
4. от 61 до 80%	7	4	0,156	0,160	0,978	1,000	0,022
5. от 81 до 100%	1	0	0,022	0	1,000	1,000	0

Максимальная выявленная между двумя накопленными эмпирическими частостями разность составляет 0,218.

Эта разность оказывается накопленной во второй категории прогноза. Попробуем использовать верхнюю границу данной категории в качестве критерия для разделения обеих выборок на подгруппу, где "есть эффект" и подгруппу, где "нет эффекта". Будем считать, что "эффект есть", если данный врач прогнозирует от 41 до 100% приемных с фондами в 1993 году, и что "эффекта нет", если данный врач прогнозирует от 0 до 40% приемных с фондами в 1993 году. Мы объединяем категории прогноза 1 и 2, с одной стороны, и категории прогноза 3, 4 и 5, с другой. Полученное распределение прогнозов представлено в Табл. 5.10.

Таблица 5.10

Распределение прогнозов у врачей с фондами и врачей без фондов

Прогнозируемая доля приемных с фондами(%1	Эмпирические частоты выбора данной категории прогноза		Суммы
	врачами с фондом (n_1=45)	врачами без фонда (n_2=25)	
1. от 0 до 40%	19	16	35
2. от 41 до 100%	26	9	35
Суммы	45	25	70

Полученную таблицу (Табл. 5.10) мы можем использовать, проверяя разные гипотезы путем сопоставления любых двух ее ячеек. Мы помним, что это так называемая четырехклеточная, или четырехпольная, таблица.

В данном случае нас интересует, действительно ли врачи, уже располагающие

фондами, прогнозируют больший размах этого движения в будущем, чем врачи, не располагающие фондами. Поэтому мы условно считаем, что "эффект есть", когда прогноз попадает в категорию от 41 до 100%. Для упрощения расчетов нам необходимо теперь повернуть таблицу на 90°, вращая ее по направлению часовой стрелки. Можно сделать это даже буквально, повернув книгу вместе с таблицей. Теперь мы можем перейти к рабочей таблице для расчета критерия φ* - углового преобразования Фишера.

Таблица 5.11

Четырехклеточная таблица для подсчета критерия φ* Фишера для выявления различий в прогнозах двух групп общепрактикующих врачей

Группа	Есть эффект -прогноз от 41 до 100%	Нет эффекта -прогноз от 0 до 40%	Всего
I группа - врачи, взявшие фонд	26 (57.8%)	19 (42.2%)	45
II группа - врачи, не взявшие фонда	9 (36.0%)	16 (64.0%)	25
Всего	35	35	70

Сформулируем гипотезы.

H_0: Доля лиц, прогнозирующих распространение фондов на 41%-100% всех врачебных приемных, в группе врачей с фондами не больше, чем в группе врачей без фондов.

H_1: Доля лиц, прогнозирующих распространение фондов на 41%-100% всех приемных, в группе врачей с фондами больше, чем в группе врачей без фондов.

Определяем величины φ_1 и φ_2 по Таблице XII приложения 1. Напомним, что φ_1 - это всегда угол, соответствующий большей процентной доле.

$\varphi_{1(57,8\%)}=1,727$

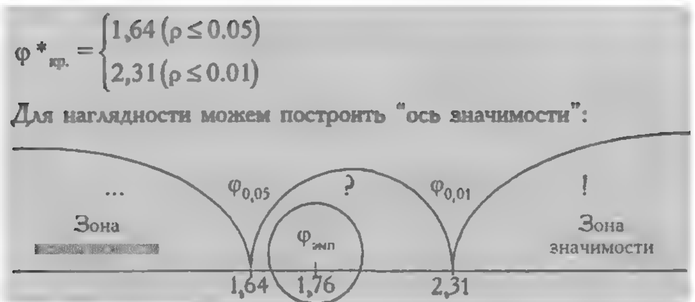
$\varphi_{2(36,0\%)}=1,287$

Теперь определим эмпирическое значение критерия φ*:

$$\varphi^* = (1,727 - 1,287) \cdot \sqrt{\frac{45 \cdot 25}{45 + 25}} = 0,440 \cdot 4,009 = 1,764$$

По Табл. XIII Приложения 1 определяем, какому уровню значимости соответствует эта величина: p=0,039.

По той же таблице Приложения 1 можно определить критические значения критерия φ*:



Ответ: Но отвергается (p=0,039). Доля лиц, прогнозирующих распространение фондов на 41-100% всех приемных, в группе врачей, взявших фонд, превышает эту долю в группе врачей, не взявших фонда.

Иными словами, врачи, уже работающие в своих приемных на отдельном бюджете, прогнозируют более широкое распространение этой практики в текущем году, чем врачи, пока еще не согласившиеся перейти на самостоятельный бюджет. Интерпретации этого результата многозначны. Например, можно предположить, что врачи каждой из групп подсознательно считают свое поведение более типичным. Это может означать также, что врачи, уже перешедшие на самостоятельный бюджет, склонны преувеличивать размах этого движения, так как им нужно оправдать свое решение. Выявленные различия могут означать и нечто такое, что вовсе выходит за рамки поставленных в исследовании вопросов. Например, что активность врачей, работающих на самостоятельном бюджете,

способствует заострению различий в позициях обеих групп. Они проявили большую активность, когда согласились взять фонды, они проявили большую активность, когда взяли на себя труд ответить на почтовый опросник; они проявляют большую активность, когда прогнозируют большую активность других врачей в получении фондов.

Так или иначе, мы можем быть уверены, что выявленный уровень статистических различий - максимально возможный для этих реальных данных. Мы установили с помощью критерия λ точку максимального расхождения между двумя распределениями и именно в этой точке разделили выборки на две части.

АЛГОРИТМ 17
Расчет критерия φ*

1. Определить те значения признака, которые будут критерием для разделения испытуемых на тех, у кого "есть эффект" и тех, у кого "нет эффекта". Если признак измерен количественно, использовать критерий λ для поиска оптимальной точки разделения.

2. Начертить четырехклеточную таблицу из двух столбцов и двух строк. Первый столбец - "есть эффект"; второй столбец - "нет эффекта"; первая строка сверху - 1 группа (выборка); вторая строка - 2 группа (выборка).

3. Подсчитать количество испытуемых в первой группе, у которых "есть эффект", и занести это число в левую верхнюю ячейку таблицы.

4. Подсчитать количество испытуемых в первой выборке, у которых "нет эффекта", и занести это число в правую верхнюю ячейку таблицы. Подсчитать сумму по двум верхним ячейкам. Она должна совпадать с количеством испытуемых в первой группе.

5. Подсчитать количество испытуемых во второй группе, у которых "есть эффект", и занести это число в левую нижнюю ячейку таблицы.

6. Подсчитать количество испытуемых во второй выборке, у которых "нет эффекта", и занести это число в правую нижнюю ячейку таблицы. Подсчитать сумму по двум нижним ячейкам. Она должна совпадать с количеством испытуемых во второй группе (выборке).

7. Определить процентные доли испытуемых, у которых "есть эффект", путем отнесения их количества к общему количеству испытуемых в данной группе (выборке). Записать полученные процентные доли соответственно в левой верхней и левой нижней ячейках таблицы в скобках, чтобы не перепутать их с абсолютными значениями.

8. Проверить, не равняется ли одна из сопоставляемых процентных долей нулю. Если это так, попробовать изменить это, сдвинув точку разделения групп в ту или иную сторону. Если это невозможно или нежелательно, отказаться от критерия φ* и использовать критерий χ^2.

9. Определить по Табл. XII Приложения 1 величины углов φ для каждой из сопоставляемых процентных долей.

10. Подсчитать эмпирическое значение φ* по формуле:

$$\varphi^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

где: φ_1 - угол, соответствующий большей процентной доле;

φ_2 - угол, соответствующий меньшей процентной доле;

n_1 - количество наблюдений в выборке 1;

n_2 - количество наблюдений в выборке 2.

11. Сопоставить полученное значение φ* с критическими значениями: $\varphi^* \leq 1{,}64$ ($p<0{,}05$) и $\varphi^* \leq 2{,}31$ ($p<0{,}01$).

Если $\varphi^*_{эмп} \leq \varphi^*_{кр}$. H_0 отвергается.

При необходимости определить точный уровень значимости полученного $\varphi^*_{эмп}$ по Табл. XIII Приложения 1.

5.3. Биномиальный критерий ш Назначение критерия m

Критерий предназначен для сопоставления частоты встречаемости какого-либо эффекта с теоретической или заданной частотой его встречаемости.

Он применяется в тех случаях, когда обследована лишь одна выборка объемом не более 300 наблюдений, в некоторых задачах - не больше 50 наблюдений.

Описание критерия

Биномиальный критерий m позволяет оценить, насколько эмпирическая частота интересующего нас эффекта *превышает* теоретическую, среднестатистическую или какую-то заданную частоту, соответствующую вероятности случайного угадывания, среднему проценту успешности в выполнении данного задания, допустимому проценту брака и т.п.

Биномиальный критерий незаменим, если налицо 2 условия:

а) обследована лишь *одна* выборка испытуемых, и нет возможности или смысла делить эту выборку на две части с целью дальнейшего применения критерия, φ*, так как для нас по каким-то причинам важно исследовать частоту встречаемости признака *в выборке в целом;*

б) в обследованной выборке *менее 30 испытуемых,* что не позволяет нам применить критерий χ^2.

Если в нашей выборке больше 30 испытуемых, мы все же можем использовать критерий m и тем самым сэкономить время на подсчете χ^2.

Эмпирическая частота наблюдений, в которых проявляется интересующий нас эффект, обозначается как т. Это и есть эмпирическое [значение критерия т. Если $m_{эмп}$ равен или превышает $m_{кр}$, то различия достоверны.

Гипотезы

H_0: Частота встречаемости данного эффекта в обследованной выборке не превышает теоретической (заданной, ожидаемой, предполагаемой).

H_1: Частота встречаемости данного эффекта в обследованной выборке превышает теоретическую (заданную, ожидаемую, предполагаемую).

Графическое представление биномиального критерия

Критерий определяет, достаточно ли эмпирическая частота встречаемости признака превышает заданную, "перевешивает" ее. Можно представить себе это как взвешивание эмпирической и теоретической частот на чашечных весах (Рис. 5.4). Весы реагируют только на такие, различия в весе, которые соответствуют по крайней мере минимальному уровню значимости p<0,05.

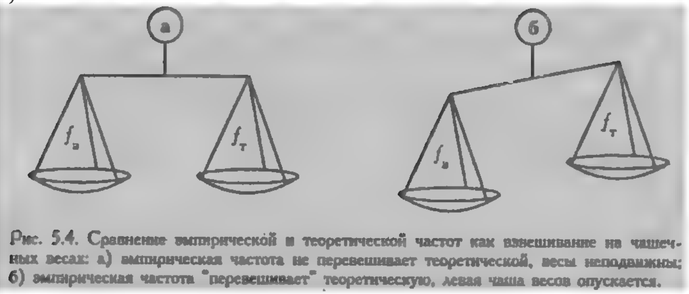


Рис. 5.4. Сравнение эмпирической и теоретической частот как взвешивание на чашечных весах: а) эмпирическая частота не перевешивает теоретической, весы неподвижны; б) эмпирическая частота "перевешивает" теоретическую, левая чаша весов опускается.

Ограничения биномиального критерия

1. В выборке должно быть не менее 5 наблюдений. В принципе возможно применение критерия и при 2≤n<5, но лишь в отношении определенного типа задач (см. Табл. XV Приложения 1).

2. Верхний предел численности выборки зависит от ограничений, определяемых пп.3-8 и варьирует в диапазоне от 50 до 300 наблюдений, что определяется имеющимися

таблицами критических значений.

3. Биномиальный критерий m позволяет проверить лишь гипотезу о том, что частота встречаемости интересующего нас эффекта в обследованной выборке *превышает* заданную вероятность Р. Заданная вероятность при этом должна быть: $P \leq 0{,}50$.

4. Если мы хотим проверить гипотезу о том, что частота встречаемости интересующего нас эффекта достоверно *ниже* заданной вероятности, то при Р=0,50 мы можем сделать это с помощью уже известного критерия знаков G, при Р>0,50 мы должны преобразовать гипотезы в противоположные, а при Р<0,50 придется использовать критерий χ^2.

По Табл. 5.12 легко определить, какой из путей для нас доступен.

Таблица 5.12

Выбор критерия для сопоставлений эмпирической частоты с теоретической при разных вероятностях исследуемого эффекта Р и разных гипотезах.

Заданные вероятности	H_1: $f_{эмп}$ достоверно выше $f_{теор}$	H_1: $f_{эмп}$ достоверно ниже $f_{теор}$
P<0,50	А m для $2 \leq n \leq 50$	Б χ^2 для $n \geq 30$
P=0,50	В m для $5 \leq n \leq 300$	Г G для $5 \leq n \leq 300$
P>0,50	Д χ^2 для $n \leq 30$	Е m для $2 \leq n \leq 50$

Пояснения к Табл. 5.12

А) Если заданная вероятность Р<0,50, а $f_{эмп}>f_{теор}$ (например, допустимый уровень брака - 15%, а в обследованной выборке получено значение в 25%), то биномиальный критерий применим для объема выборки $2 \leq n \leq 50$.

Б) Если заданная вероятность Р<0,50, а $f_{эмп}>f_{теор}$ (например, допустимый уровень брака - 15%, а в обследованной выборке наблюдается 5% брака), то биномиальный критерий неприменим и следует применять критерий χ^2 (см. Пример 2).

В) Если заданная вероятность Р=0,50, а $f_{эмп}>f_{теор}$ (например, вероятность выбора каждой из равновероятных альтернатив Р=0,50, а в обследованной выборке одна из альтернатив выбирается чаще, чем **в** половине случаев), то биномиальный критерий применим для объема выборки $5 \leq n \leq 300$.

Г) Если заданная вероятность Р=0,50, а $f_{эмп}>f_{теор}$ (например, вероятность выбора каждой из равновероятных альтернатив Р=0,50, а в обследованной выборке одна из альтернатив наблюдается реже, чем в половине случаев), то вместо биномиального критерия применяется критерий знаков G, являющийся "зеркальным отражением" биномиального критерия при Р=0,50. Допустимый объем выборки: $5 \leq n \leq 300$.

Д) Если заданная вероятность Р>0,50, а $f_{эмп}>f_{теор}$ (например, среднестатистический процент решения задачи - 80%, а в обследованной выборке он составляет 95%), то биномиальный критерий неприменим и следует применять критерий χ^2 (см. Пример 3).

Е) Если заданная вероятность Р>0,50, а $f_{эмп}>f_{теор}$ (например, среднестатистический процент решения задачи - 80%, а в обследованной выборке он составляет 60%), то биномиальный критерий применим при условии, что в качестве "эффекта" мы будем рассматривать более редкое событие - неудачу в решении задачи, вероятность которого Q=l—Р=1—0,80=0,20 и процент встречаемости в данной выборке: 100%—75%=25%. Эти преобразования фактически сведут данную задачу к задаче, предусмотренной п. А. Допустимый объем выборки: $2 \leq n \leq 50$ (см. пример 3).

Пример 1

В процессе тренинга сенситивности в группе из 14 человек выполнялось упражнение "Психологический прогноз". Все участники должны были пристально вглядеться в одного и того же человека, который сам пожелал быть испытуемым в этом упражнении. Затем каждый из участников задавал испытуемому вопрос, предполагавший два заданных варианта ответа, например: "Что в тебе преобладает: отстраненная наблюдательность или включенная эмпатия?" "Продолжал бы ты работать или нет, если бы у тебя появилась материальная возможность не работать?" "Кто тебя больше утомляет - люди нахальные или занудные?" и т. п. Испытуемый должен был лишь молча выслушать вопрос, ничего не

отвечая. Во время этой паузы участники пытались определить, как он ответит на данный вопрос, и записывали свои прогнозы. Затем ведущий предлагал испытуемому дать ответ на заданный вопрос. Теперь каждый участник мог определить, совпал ли его прогноз с ответом испытуемого или нет. После того, как было задано 14 вопросов (13 участников + ведущий), каждый сообщил, сколько у него получилось точных прогнозов. В среднем было по 7-8 совпадений, но у одного из участников их было 12, и группа ему спонтанно зааплодировала. У другого участника, однако, оказалось всего 4 совпадения, и он был очень этим огорчен.

Имела ли группа статистические основания для аплодисментов?

Имел ли огорченный участник статистические основания для грусти?

Начнем с первого вопроса.

По-видимому, группа будет иметь статистические основания для аплодисментов, если частота правильных прогнозов у участника А превысит теоретическую частоту случайных угадываний. Если бы участник прогнозировал ответ испытуемого случайным образом, то, в соответствии с теорией вероятностей, шансы случайно угадать или не угадать ответ на данный вопрос у него были бы равны P=Q=0,5. Определим теоретическую частоту правильных случайных угадываний:

$f_{теор}=n \cdot P$

где n - количество прогнозов;

P - вероятность правильного прогноза при случайном угадывании.

$f_{теор}=14 \cdot 0,5=7$

Итак, нам нужно определить, "перевешивают" ли 12 реально данных правильных прогнозов 7 правильных прогнозов, которые могли бы быть у данного участника, если бы он прогнозировал ответ испытуемого случайным образом.

Требования, предусмотренные ограничением 3, соблюдены: P=0,50; $f_{эмп}>f_{теор}$. Данный случай относится к варианту "В" Табл. 5.12.

Мы можем сформулировать гипотезы.

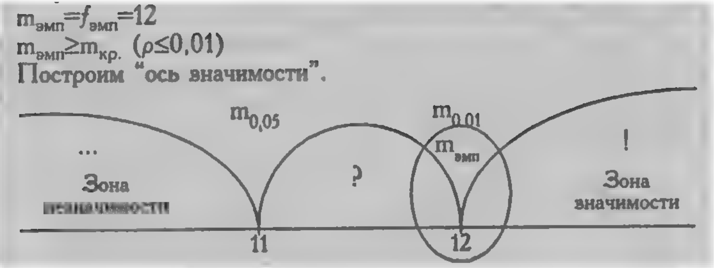
H_0: Количество точных прогнозов у участника А не превышает частоты, соответствующей вероятности случайного угадывания.

H_1: Количество точных прогнозов у участника А превышает частоту, соответствующую вероятности случайного угадывания.

По Табл. XIV Приложения 1 определяем критические значения критерия m при n=14, P=0,50:

$$m_{кр.}=\begin{cases}11 \ (\rho \le 0,05)\\ 12 \ (\rho \le 0,01)\end{cases}$$

Мы помним, что за эмпирическое значение критерия m принимается эмпирическая частота:



Зона значимости простирается вправо, в область более высоких значений m (более "весомых", если использовать аналогию с весами), а зона незначимости - в область более низких, "невесомых", значений m.

Ответ: H_0 отвергается. Принимается H_1. Количество точных прогнозов у участника А превышает (или по крайней мере равняется) критической частоте вероятности случайного угадывания ($p \le 0,01$). Группа вполне обоснованно ему аплодировала!

Теперь попробуем ответить на второй вопрос задачи.

По-видимому, основания для грусти могут появиться, если количество правильных

прогнозов оказывается достоверно ниже теоретической частоты случайных угадываний. Мы должны определить, 4 точных прогноза участника Б - это достоверно меньше, чем 7 теоретически возможных правильных прогнозов при случайном угадывании или нет?

В данном случае P=0,50; $f_{эмп}>f_{теор}$. В соответствии с ограничением 4, в данном случае мы должны применить критерий знаков, который по существу является зеркальным отражением или "второй стороной" одностороннего биномиального критерия (вариант "Г" Табл. 5.12). Вначале нам нужно определить, что является типичным событием для участника Б. Это неправильные прогнозы, их 10. Теперь мы определяем, достаточно ли мало у него нетипичных правильных прогнозов, чтобы считать перевешивание неправильных прогнозов достоверным.

Сформулируем гипотезы.

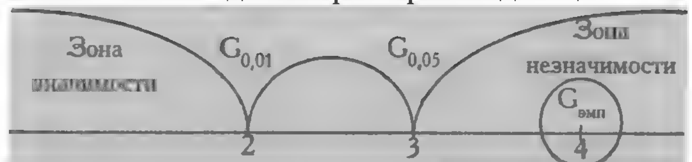
H_0: Преобладание неправильных прогнозов у участника Б является случайным.

H_1: Преобладание неправильных прогнозов у участника Б не является случайным.

По Табл. V Приложения 1 определяем критические значения критерия знаков G для n=14:

$$G_{кр} = \begin{cases} 3\,(p \le 0{,}05) \\ 2\,(p \le 0{,}01) \end{cases}$$

Построим "ось значимости". Мы помним, что в критерии знаков зона значимости находится слева, а зона незначимости - справа, так как чем меньше нетипичных событий, тем типичные события являются более достоверно преобладающими.



Эмпирическое значение критерия G определяется как количество нетипичных событий. В данном случае:

$G_{эмп}=4$

$G_{эмп}>G_{теор}$

Эмпирическое значение критерия G попадает в зону незначимости.

Ответ: H_0 принимается. Преобладание неправильных прогнозов у участника Б является случайным.

Участник Б не имел достаточных статистических оснований для огорчения. Дело, однако, в том, что психологическая "весомость" отклонения его оценки значительно перевешивает статистическую. Всякий практикующий психолог согласится, что повод для огорчения у участника Б все же был.

Важная особенность биномиального критерия и критерия знаков состоит в том, что они превращают уникальность, единственность и жизненную резкость произошедшего события в нечто неотличимое от безликой и всепоглощающей случайности. Учитывая это, лучше использовать биномиальный критерий для решения более отвлеченных, формализованных задач, например, для уравновешивания выборок по признаку пола, возраста, профессиональной принадлежности и т. п.

При оценке же личностно значимых событий оказывается, что статистическая сторона дела не совпадает с психологической больше, чем при использовании любого из других критериев.

Пример 2

В тренинге профессиональных наблюдателей допускается, чтобы наблюдатель ошибался в оценке возраста ребенка не более чем на 1 год в ту или иную сторону. Наблюдатель допускается к работе, если он совершает не более 15% ошибок, превышающих отклонение на 1 год. Наблюдатель Н допустил 1 ошибку в 50-ти попытках, а наблюдатель К - 15 ошибок в 50-ти попытках. Достоверно ли отличаются эти результаты от контрольной величины?

Определим частоту допустимых ошибок при $n = 50$:

$f_{теор}=n \cdot P=50 \cdot 0{,}15=7{,}5$

Для наблюдателя Н $f_{эмп}<f_{теор}$. Для наблюдателя К $f_{эмп}>f_{теор}$

Сформулируем гипотезы для наблюдателя Н.

H_0: Количество ошибок у наблюдателя Н не меньше, чем это предусмотрено заданной величиной.

H_1: Количество ошибок у наблюдателя Н меньше, чем это предусмотрено заданной величиной.

В данном случае P=0,15<0,50; $f_{эмп}>f_{теор}$.

Этот случай попадает под вариант Б Табл. 5. 12. Нам придется применить критерий *у},* сопоставляя полученные эмпирические частоты ошибочных и правильных ответов с теоретическими частотами, составляющими, соответственно, 7,5 для ошибочного ответа и (50-7,5)=42,5 для правильного ответа. Подсчитаем χ^2 по формуле, включающей поправку на непрерывность[24]:

$$\chi^2 = \sum \frac{\left(\left|f_j^э - f_j^т\right| - 0{,}5\right)^2}{f_j^т} = \frac{\left(\left|1-7{,}5\right|-0{,}5\right)^2}{7{,}5} + \frac{\left(\left|49-42{,}5\right|-0{,}5\right)^2}{42{,}5} =$$

$$= 4{,}800 + 0{,}85 = 5{,}65$$

По Табл. IX Приложения 1 определяем критические значения χ^2 при v=1:

$$\chi^2_{кр.} = \begin{cases} 3{,}841 \ (\rho \le 0{,}05) \\ 6{,}635 \ (\rho \le 0{,}01) \end{cases}$$

$$\chi^2_{эмп} > \chi^2_{кр.} \ (\rho \le 0{,}05)$$

Ответ: H_0 отвергается. Количество ошибок у наблюдателя Н меньше, чем это предусмотрено заданной величиной ($p \le 0{,}05$)

Сформулируем гипотезы для наблюдателя К.

H_0: Количество ошибок у наблюдателя К не больше, чем это предусмотрено заданной величиной.

H_1: Количество ошибок у наблюдателя К больше, чем это предусмотрено заданной величиной.

В данном случае P=0,15<0,5; $f_{эмп}>f_{теор}$. Этот случай подпадает под вариант А Табл. 5.12. Мы можем применить биномиальный критерий, поскольку n=50.

По Табл. XV Приложения 1 определяем критические значения при п=50, P=0,15, Q=0,85:

$$m_{кр} = \begin{cases} 13 \ (p \le 0{,}05) \\ 16 \ (p \le 0{,}01) \end{cases}$$

$$m_{эмп} = f_{эмп} = 15$$

$$m_{эмп} > m_{кр} \ (\rho \le 0{,}05)$$

Ответ: H_0 отвергается. Количество ошибок у наблюдателя Н меньше, чем это предусмотрено заданной величиной (p<0,05).

Пример 3

В примере 1 параграфа 5.2 мы сравнивали процент справившихся с экспериментальной задачей испытуемых в двух группах. Теперь мы можем сопоставить процент успешности каждой группы со среднестатистическим процентом успешности. Данные представлены в Табл. 5.13.

Таблица 5.13

Показатели успешности решения задачи в двух группах испытуемых

	Количество испытуемых, решивших задачу		Количество испытуемых, не решивших задачу		Суммы
1 группа (n_1=20)	12	(60%)	8	(40%)	20
2 группа (п7=25)	10	(40%)	15	(60%)	25

[24] Поправка на непрерывность вносится во всех случаях, когда признак принимает всего два значения и число степеней свободы поэтому равно 1 (см. параграф 4.2)

Суммы	22	23	45

Среднестатистический показатель успешности в решении этой задачи - 55%. Определим теоретическую частоту правильных ответов для групп 1 и 2:

$f_{теор\ 1}=n_1 \cdot P=20 \cdot 0,55=11,00$

$f_{теор\ 2}=n_2 \cdot P=25 \cdot 0,55=13,75$

Для группы 1, следовательно, P=0,55>0,50; $f_{эмп}=12>f_{теор}$. Этот случай соответствует варианту "Д" Табл. 5.12. Мы должны были бы применить критерий χ^2, но у нас всего 20 наблюдений: n<30. Ни биномиальный критерий, ни критерий χ^2 неприменимы. Остается критерий φ* Фишера, который мы сможем применить, если узнаем, сколько испытуемых было в выборке, по которой определялся среднестатистический процент.

Далее, для группы 2: P=0,55>0,50; $f_{эмп}=10>f_{теор}$. Этот случай соответствует варианту "Е" Табл. 5.12. Мы можем применить биномиальный критерий, если будем считать "эффектом" неудачу в решении задачи. Вероятность неудачи Q=l—P=1—0,55=0,45. Новая эмпирическая частота составит: $f_{эмп}=25-10=15$.

Сформулируем гипотезы.

H_0: Процент неудач в обследованной выборке не превышает заданного процента неудач.

H_1 Процент неудач в обследованной выборке превышает заданный процент неудач.

По Табл. XV Приложения 1 определяем критические значения для n=25, P=0,45, Q=0.55 (мы помним, что P и Q поменялись местами):

$$m_{кр}=\begin{cases}16\ (p \le 0,05)\\18\ (p \le 0,01)\end{cases}$$

$m_{эмп}=f_{эмп}=15$

$m_{эмп}<m_{кр}$

Ответ: H_0 принимается. Процент неудач в обследованной выборке не превышает заданного процента неудач.

Сформулируем общий алгоритм применения критерия m.

АЛГОРИТМ 18

Применение биномиального критерия m

1. Определить теоретическую частоту встречаемости эффекта по формуле:

$F_{теор}=n \cdot P$.

где n - количество наблюдений в обследованной выборке;

P - заданная вероятность исследуемого эффекта.

По соотношению эмпирической и теоретической частот и заданной вероятности P определить, к какой ячейке Табл. 5.12 относится данный случай сопоставлений.

Если биномиальный критерий оказывается неприменимым, использовать тот критерий, который указан в соответствующей ячейке Табл. 5.12

2. Если критерий m применим, то определить критические значения m по Табл. XVI (при P=0,50) или по табл. XV (при P<0,50) для данных n и P.

3. Считать $m_{эмп}$ эмпирическую частоту встречаемости эффекта в обследованной выборке: $m_{эмп}=f_{эмп}$.

4. Если $m_{эмп}$ превышает критические значения, это означает, что эмпирическая частота достоверно превышает частоту, соответствующую заданной вероятности.

5.4. Многофункциональные критерии как эффективные заменители традиционных критериев

Как было показано в предыдущих параграфах, многофункциональные критерии, главным образом критерий φ*, применим к решению всех трех типов задач, рассмотренных в Главах 2-4: сопоставление уровней, определение сдвигов и сравнение распределений признака. В тех случаях, когда обследованы две выборки испытуемых, критерий φ* может эффективно заменять или, по крайней мере, эффективно дополнять традиционные критерии: Q - критерий Розенбаума, U - критерий Манна-Уитни, критерий

χ^2 Пирсона и критерий λ Колмогорова-Смирнова.

В особенности полезна такая замена в следующих случаях:

Случай 1. Другие критерии неприменимы

Часто бывает так, что критерий Q неприменим вследствие совпадения диапазонов двух выборок, а критерий U неприменим вследствие того, что количество наблюдений n>60.

В качестве примера сошлемся на задачу сравнения сдвигов оценок в экспериментальной и контрольной группах после просмотра видеозаписи и чтения текста о пользе телесных наказаний (см. параграф 3.2). Сдвиги в двух группах являются показателями, полученными независимо в двух группах испытуемых. Задача сравнения таких показателей сдвига - это частный случай задачи сопоставления двух групп по уровню значений какого-либо признака. Такие задачи решаются с помощью критериев Q Розенбаума и U Манна-Уитни (см. Табл. 3.1). Сводные данные по сдвигам в двух группах представлены в Табл. 5.14.

Таблица 5.14

Эмпирические частоты сдвигов разной интенсивности и направления в экспериментальной и контрольной группах после предъявления видеозаписи или письменного текста

Значения сдвига	Количество сдвигов в экспериментальной группе (n_1=16)	Количество сдвигов в контрольной группе (n_2=23)	Суммы
+5	0	1	1
+2	3	5	8
+1	19	11	30
0	38	65	103
-1	4	8	12
-2	0	2	2
Суммы	64	92	156

В экспериментальной группе значения сдвигов варьируют от —2 до +2, а в контрольной группе от —2 до +5. Критерий Q неприменим. Критерий U неприменим, поскольку количество наблюдений (сдвигов) в каждой группе больше 60.

Применяем критерий φ*. Построим вначале четырехклеточную таблицу для положительных сдвигов, а затем - для нулевых.

Таблица 5.15

Четырехклеточная таблица для подсчета критерия φ* при сопоставлении долей положительных сдвигов в экспериментальной и контрольной группах

Группы	Есть эффект : сдвиг положительный	"Нет эффекта".' сдвиг отрицательный или нулевок	Суммы
Группа 1 экспериментальная	22 (34,4%)	42 (656%)	64
Группа 2 контрольная	17 (18,5%)	75 (81,5%)	92
Суммы	39	117	156

Сформулируем гипотезы.

H_0: Доля положительных сдвигов в экспериментальной группе не больше, чем в контрольной.

H_1: Доля положительных сдвигов в экспериментальной группе больше, чем в контрольной.

Далее действуем по Алгоритму 17.

$\varphi_{1\ (34,4\%)}=1,254$

$\varphi_{2\ (18,5\%)}=0,889$

$$\varphi^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (1,254 - 0,889) \cdot \sqrt{\frac{64 \cdot 92}{64 + 92}} =$$

$$= 0,365 \cdot \sqrt{37,74} = 0,365 \cdot 6,143 = 2,242$$

$$\varphi^*_{кр} = \begin{cases} 1,64\ (p \le 0.05) \\ 2,31\ (p \le 0.01) \end{cases}$$

Мы можем и точно определить уровень статистической значимости полученного результата по Табл. XIII Приложения 1:

при $\varphi^*_{эмп}=2,242$ $p=0,013$.

Ответ: H_0 отклоняется. Принимается H_1. Доля положительных сдвигов в экспериментальной группе больше, чем в контрольной ($p<0,013$).

Теперь перейдем к вопросу о меньшей доле нулевых сдвигов в экспериментальной группе.

Таблица 5.16

Четырехклеточная таблица для подсчета критерия φ* при сопоставлении долей нулевых сдвигов в экспериментальной и контрольной группах

Группы	"Есть эффект": сдвиг равен 0	"Нет сдвиг а": эффект не равен 0	Суммы
Группа 1 экспериментальная	38 (59,4%)	26 (40,6%)	64
Группа 2 контрольная	65 (70,7%)	27 (29,3%)	92
Суммы	103	53	156

Сформулируем гипотезы.

H_0: Доля нулевых сдвигов в контрольной группе не больше, чем в экспериментальной.

H_1: Доля нулевых сдвигов в контрольной группе больше, чем в экспериментальной. Далее действуем по Алгоритму 17.

$\varphi_{1\ (70,7\%)}=1,998$

$\varphi_{2\ (59,4\%)}=1,760$

$$\varphi^*_{эмп} = (1,998 - 1,760) \cdot \sqrt{\frac{92 \cdot 64}{92 + 64}} = 0,238 \cdot 6,143 = 1,462$$

$\varphi^*_{эмп}<\varphi^*_{кр.}$

Ответ: H_0 принимается. Доля нулевых сдвигов в контрольной группе не больше, чем в экспериментальной.

Итак, доля положительных сдвигов в экспериментальной группе больше, но доля нулевых сдвигов - примерно такая же, как и в контрольной группе. Отметим, что в критерии знаков G все нулевые сдвиги были исключены из рассмотрения, поэтому полученный результат дает дополнительную информацию, которую не мог дать критерий знаков.

Случай 2. Другие критерии неэффективны или слишком громоздки

В качестве примера можно указать на задачу с сопоставлением показателей недостаточности в группах с большей и меньшей энергией вытеснения (см. Табл. 5.4).

Критерий Q дает незначимый результат:

$Q=S_1+S_2=4+0=4$

$$\text{при } Q_{кр.} = \begin{cases} 8\ (\rho \le 0,05) \\ 9\ (\rho \le 0,01) \end{cases}$$

Критерий U в данном случае применим и даже дает значимый результат ($U_{эмп}=154,5$; $p<0,05$), однако ранжирование показателей, многие из которых имеют одно и то же

значение (например, значение 30 баллов встречается 13 раз), представляет определенные трудности.

Как мы помним, с помощью критерия φ* удалось доказать, что наиболее высокие показатели недостаточности (30 и более баллов) встречаются в группе с большей энергией вытеснения чаще, чем в группе с меньшей энергией вытеснения (p=0,008) и что, с другой стороны, самые низкие (нулевые) показатели встречаются чаще также в этой группе ($p \leq 0{,}05$).

Другим примером может служить задача сопоставления распределения выборов желтого цвета в отечественной выборке и в выборке Х.Клара (см. параграф 4.3).

Критерий λ не выявил достоверных различий между двумя распределениями, однако позволил нам установить точку максимального накопленного расхождения между ними. Из Табл. 4.19 следует, что такой точкой является вторая позиция желтого цвета. Построим четы-рехклеточную таблицу, где "эффектом" будет считаться попадание желтого цвета на одну из первых двух позиций.

Таблица 5.17

Четырехклеточная таблица для расчета φ* при сопоставлении отечественной выборки (n_1=102) и выборки Х.Клара (n_2=800) по положению желтого цвета в ряду предпочтений

Выборки	"Есть эффект": желтый цвет на первых двух позициях	"Нет эффекта": желтый цвет на позициях 3-8	Суммы
Выборка 1 - отечественная	39 (38.2%)	63 (61,8%)	102
Выборка 2 - Х.Клара	211 (26,4%)	589 (73,6%)	800
Суммы	250	652	902

Сформулируем гипотезы:

H_0: Доля лиц, помещающих желтый цвет на одну из первых двух позиций, в отечественной выборке не больше, чем в выборке Х.Клара.

H_1: Доля лиц, поместивших желтый цвет на одну из первых двух позиций, в отечественной выборке больше, чем в выборке Х. Клара.

Далее действуем по Алгоритму 17.

$\varphi_{1\,(38,2\%)}=1{,}333$

$\varphi_{2\,(26,4\%)}=1{,}079$

$$\varphi^*_{эмп} = (1{,}333 - 1{,}079) \cdot \sqrt{\frac{102 \cdot 800}{102 + 800}} = 0{,}254 \cdot \sqrt{90{,}47} =$$

$$= 0{,}254 \cdot 9{,}512 = 2{,}416$$

Как мы помним,

$$\varphi^*_{кр.} = \begin{cases} 1{,}64\ (\rho \leq 0{,}05) \\ 2{,}31\ (\rho \leq 0{,}01) \end{cases}$$

$\varphi^*_{эмп} > \varphi^*_{кр.}$

Ответ: H_0 отклоняется. Принимается H_1: *Доля* лиц, поместивших желтый цвет на одну из первых двух позиций, в отечественной выборке больше, чем в выборке Х.Клара ($p<0{,}01$).

Мы еще раз столкнулись с тем случаем, когда критерий A, сам по себе не выявляет достоверных различий, но помогает максимально использовать возможности критерия φ*.

Случай 3. Другие критерии слишком трудоемки

Этот случай чаще всего относится к критерию χ^2. Заменить его критерием φ* можно при условии, если сравниваются распределения признака в двух выборках, а сам признак принимает всего два значения[25].

В качестве примера можно привести задачу с соотношением мужских и женских

---

[25] В принципе признак может принимать и большее количество значений, так как любую шкалу, как мы убедились, можно свести к альтернативной шкале "Есть эффект" - "Нет эффекта".

имен в записных книжках двух психологов (см. п. 4.2, Табл. 4.11).

Преобразуем Табл. 4.11 в четырехклеточную таблицу, где "эффектом" будем считать мужские имена.

Таблица 5.18

Четырехклеточная таблица для подсчета φ* при сопоставлении записных книжек двух психологов по соотношению мужских и женских имен

Группы	"Есть аффект": мужские имена	"Нет аффекта»: женские имена	Суммы
Группа 1 - выборка имен в книжке Х.	22 (32,8%)	45 (67,2%)	67
Группа 2 - выборка имен в книжке С.	59 (35,1%)	109 (64,9%)	168
Суммы	81	154	235

Сформулируем гипотезы.

H_0: Доля мужских имен в записной книжке С. не больше, чем в записной книжке Х.

H_1: Доля мужских имен в записной книжке С. больше, чем в записной книжке Х.

Далее действуем по алгоритму.

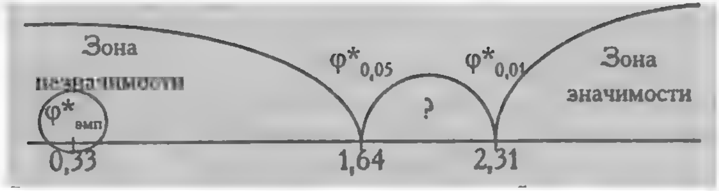
$\varphi_{2\,(35,1\%)}=1,268$

$\varphi_{2\,(32,8\%)}=1,220$

$$\varphi^*_{эмп} = (1,268-1,220)\cdot\sqrt{\frac{67\cdot 168}{67+168}} = 0,048\cdot\sqrt{47,9} =$$

$$= 0,048\cdot 6,92 = 0,332$$

По Табл. XIII Приложения 1 определяем, какому уровню достоверности соответствует это значение. Мы видим, что такого значения вообще нет в таблице. Построим "ось значимости".



Полученное эмпирическое значение - далеко в "зоне незначимости".

$f^*_{эмп} > f^*_{теор}$

Ответ: H_0 принимается. Доля мужских имен в записной книжке психолога С. не больше, чем в записной книжке психолога Х.

Исследователь сам может решить для себя, какой метод ему в данном случае удобнее применить - χ^2 или φ*. Похоже, что во втором случае меньше расчетов, хотя чуда не произошло: различия по-прежнему недостоверны.

Итак, мы убедились, что критерий φ* Фишера может эффективно заменять традиционные критерии в тех случаях, когда их применение невозможно, неэффективно или неудобно по каким-то причинам.

Биномиальный критерий m может служить заменой критерия χ^2 в случае альтернативных распределений или в случае, когда признак может принимать одно из нескольких значений и вероятность того, что он примет определенное значение, известна.

В качестве примера можно привести исследование, посвященное распределению предпочтений по 4-м типам мужественности (см. Задачу 3 к Главе 4). Если бы для испытуемых все 4 типа мужественности были одинаково привлекательными, то на первом месте примерно одинаковое количество раз оказывался бы каждый из типов. Иными словами, вероятность оказаться на первом месте для каждого типа составляла бы $^1/_4$ т.е. P=0.25.

В действительности же Национальный тип оказался на 1-м месте 19 раз, Современный - 7 раз, Религиозный - 3 раза и Мифологический - 2 раза. Можно попытаться определить, достоверно ли Национальный тип чаще оказывается на 1-м месте, чем это

предписывается вероятностью P=0,25?

Сформулируем гипотезы.

H_0; Частота попадания Национального типа мужественности на 1-е место в ряду предпочтений не превышает частоты, соответствующей вероятности P=0,25.

H_1: Частота попадания Национального типа мужественности на 1-е место в ряду предпочтений превышает частоту, соответствующую вероятности P=0,25.

Определим теоретическую частоту попадания того или иного типа мужественности на 1-е место при равновероятном выборе:

$f_{теор}=n \cdot P=31\text{-}0,25=7,75$

В данном случае соблюдаются требования, предусмотренные ограничением 3: P=0,25<0,50; $f_{эмп}>f_{теор}$. Мы можем использовать биномиальный критерий при n<50. В данном случае n=31. По Табл. XV Приложения 1 определяем критические значения m при n=31, P=0,25; Q=0,75:

$$m_{кр}=\begin{cases}13\,(p\leq 0{,}05)\\ 15\,(p\leq 0{,}01)\end{cases}$$

$m_{эмп}=f_{эмп}=19$

$m_{эмп}>m_{кр.}$

Ответ: H_0 отвергается. Частота попадания Национального типа мужественности на 1-е место в ряду предпочтений превышает частоту, соответствующую вероятности P=0,25 (p<0,01).

Итак, Национальный тип мужественности действительно чаще оказывается на 1-м месте, чем это происходило бы в том случае, если бы он выбирался на 1-е место равновероятно с другими типами.

Отметим, что мы проверяли гипотезу не об отличии данного типа мужественности от других типов, а об отличии частоты его встречаемости от теоретически возможной величины при равновероятном выборе. Все остальные типы и остальные позиции выбора остаются "за кадром" нашего рассмотрения.

Аналогичным образом можно сопоставить с теоретической частотой эмпирическую частоту попадания любого другого типа на любую другую позицию.

5.5. Задачи для самостоятельной работы

ВНИМАНИЕ!
При выборе метода решения задачи рекомендуется использовать АЛГОРИТМ 19

Задача 9

В выборке студентов факультета психологии Санкт-Петербургского университета с помощью известного "карандашного" теста определялось преобладание правого или левого глаза в прицельной, способности глаз. Совпадают ли эти данные с результатами обследования 100 студентов медицинских специальностей, представленными Т.А. Доброхотовой и Н.Н. Брагиной (1994)?

Таблица 5.19

Показатели преобладания правого и левого глаза в выборке студентов-психологов (n_1=14) и студентов-медиков (n_2=100)

	Количество испытуемых с преобладанием левого глаза	Количество испытуемых с преобладанием правого глаз;	Суммы
1. Студенты-психологи	е	8	14
2. Студенты-медики	19	81	100
Суммы	25	89	114

Задача 10

В исследовании А. А. Кузнецова (1991) изучались различия в реагировании на

вербальную агрессию между милиционерами патрульно-постовой службы и обычными гражданами. Экспериментатор в дневное время поджидал на достаточно многолюдной остановке вблизи от милицейского общежития появления мужчины в возрасте 25-35 лет и, установив с ним контакт глаз, обращался к нему с агрессивной формулой: "Ну, чего уставился?! Чего надо?!" Реакция испытуемого наблюдалась и запоминалась экспериментатором. После этого испытуемому приносились извинения и предъявлялась справка о том, что ее предъявитель является исполнителем научного эксперимента по исследованию стилей реагирования на агрессию на факультете психологии Санкт-Петербургского университета. Кроме того, экспериментатор выяснял, является ли испытуемый милиционером патрульно-постовой службы или обычным гражданином. Таким образом была собрана выборка из 25 милиционеров, которые в данный момент были не в форме и не на посту, то есть были такими же участниками гражданской жизни, как и другие граждане, и выборка из 25 граждан, не являвшихся милиционерами. Из 25 милиционеров 10 не продолжили разговора с агрессором, а 15 продолжили его, обратившись к нему с ответной фразой. Из этих 15 реакций 10 были неагрессивными и примирительными, например, "Так просто... Закурить не найдется?" или "Сколько времени, не скажешь?" или дружески: "Ух ты какой!" или мягко: "А чего ты тут стоишь?" 5 реакций были агрессивными, например, "Что?! А ну, повтори!" или "Ты что-то вякнул или мне послышалось?" или "Я тебе сейчас уставлюсь. Ну-ка, иди сюда!"

Из 25 гражданских лиц 18 предпочли не вступать в разговор, 3 человека продолжили контакт, обратившись к экспериментатору с неагрессивной, примирительной фразой вроде: "Ничего, просто смотрю" или "А может быть, вы мне понравились". Оставшиеся 4 человека продолжили контакт, дав агрессивный ответ, например, "А ты что, резкий, что ли?" и т.п.

Вопросы:

1. Можно ли утверждать, что милиционеры патрульно-постовой службы в большей степени склонны продолжать разговор с агрессором, чем другие граждане?

2. Можно ли утверждать, что милиционеры склонны отвечать агрессору более примирительно, чем гражданские лица?

Задача 11

В анкетном опросе английских общепрактикующих врачей (Курочкин М. А., Сидоренко Е. В., Чураков Ю. А., 1992) было установлено, что врачи, уже перешедшие на самостоятельный бюджет, как правило, работают в приемных с большим количеством партнеров, чем врачи, не перешедшие на самостоятельный бюджет. Возможно, врачам легче решиться взять фонды, когда их "команда" больше, но может быть, "команда" становится больше уже после того, как врачи данной приемной согласились взять фонды. Причину и следствие установить трудно. Пока необходимо установить другое: действительно ли в приемных с фондами работают большие по составу команды врачей, чем в приемных без фондов? Может ли некая фармацевтическая фирма ориентироваться на эту тенденцию при построении стратегии продвижения своего товара?

Таблица 5.20

Показатели количества партнеров у врачей с фондами и врачей без фондов (по данным М.А. Курочкина, Е.В. Сидоренко, Ю.А. Чуракова, 1992)

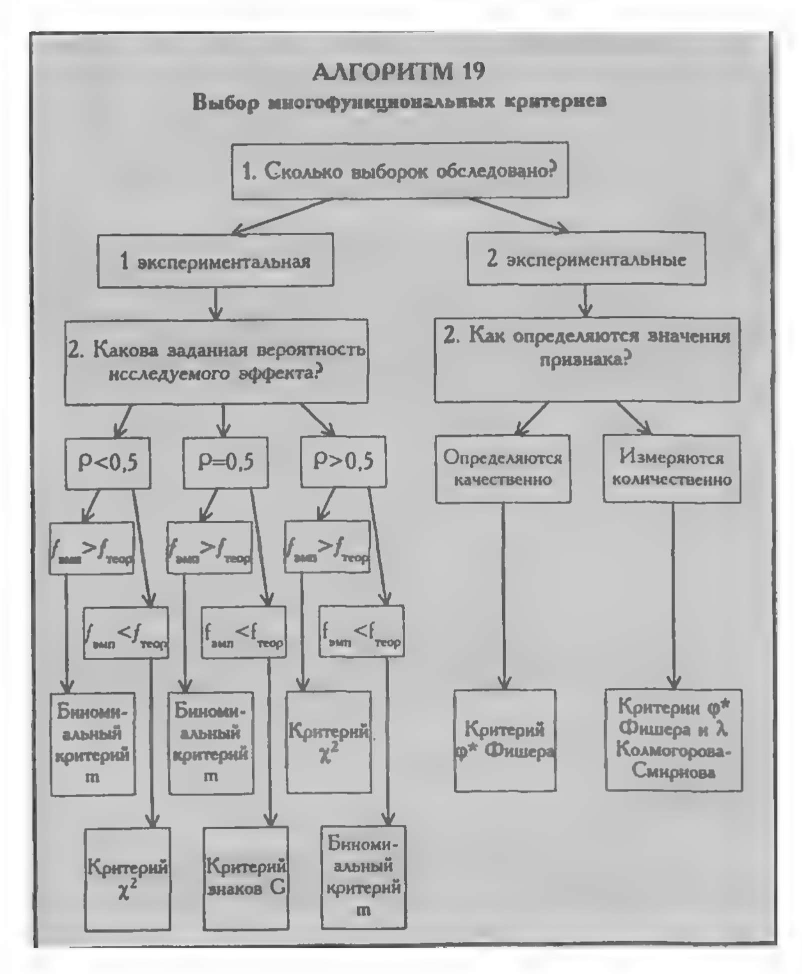
Количество партнеров		Эмпирические частоты		Всего
		в выборке врачей с фондам» (n_1=49)	в выборке врачей без фондов (n_2=28)	
1	2 и менее	2	15	17
2	3-4 партнера	6	5	11
3	5-6 партнеров	27	8	35
4	7 и более	14	0	14
Суммы		49	28	77

Задача 12

Наблюдателем установлено, что 51 человек из 70-ти выбрал правую дорожку при переходе из точки А в точку Б, а 19 человек - левую (см. параграф 4.2).

Можно ли утверждать, что правая дорожка предпочиталась достоверно чаще?

5.6. Алгоритм выбора многофункциональных критериев



5.7. Математическое сопровождение к описанию критерия φ* Фишера

Угловое преобразование позволяет перевести процентные доли, которые сами по себе имеют распределение, далекое от нормального, в величину φ, распределение которой близко к нормальному (Гублер Е.В., 1978, с. 84). Это дает определенные преимущества в том случае, если мы хотим использовать параметрические критерии, требующие нормальности распределений.

Как видно из графика на Рис. 5.1, φ нарастает в общем пропорционально процентной доле, но при этом на крайних значениях φ кривая характеризуется большей крутизной.

Благодаря этому для малых долей (меньше 20%) и больших долей (больше 80%) определение достоверности разности долей по соответствующим углам φ дает более правильные результаты, а для долей в пределах от 20 до 80% замена их углами φ дает такие же результаты, какие получаются и без этой замены, но техника вычислений при этом упрощается (Плохинский Н.А., 1970, с. 143).

Углы φ измеряются в радианах. Радиан - это угол, являющийся центральным для дуги, длина которой равна радиусу окружности (Рис. 5.5).

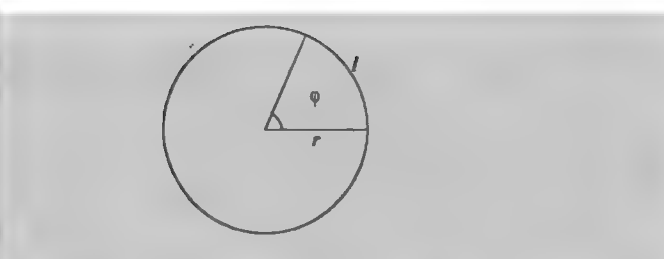


Рис. 5.5. Графическое представление центрального угла φ, величина которого измеряется отношением длины дуги, на которую этот угол опирается, к длине радиуса окружности

1 радиан равен 57°17'44".

Величина φ определяется по формуле:

$$\varphi = 2 \cdot \arcsin \sqrt{P}$$

где P - доля, выраженная в долях единицы;

arcsin - обратная синусу тригонометрическая функция.

Иными словами, синус угла φ/2 равен корню квадратному из P. Напомним, что $\sin\varphi = {}^{a}/_{c}$ (см. Рис. 5.6), а $\arcsin {}^{a}/_{c} = \varphi$

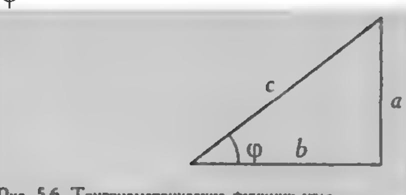


Рис. 5.6. Тригонометрические функции угла

Величину φ можно вычислить в радианах или определить по специальной таблице (Табл. XII Приложения 1).

Н.А. Плохинский использует иную формулу определения ф:

$$F_d = (\varphi_1 - \varphi_2)^2 \cdot \frac{n_1 \cdot n_2}{n_1 + n_2}$$

где φ_1 - значение угла для первой доли;

φ_2- значение угла для второй доли;

n_1 - количество наблюдений в первой выборке;

n_2 - количество наблюдений во второй выборке.

Эмпирические значения F_d сопоставляются с критическими значениями критерия F Фишера, которые определяются по таблице для степеней свободы v_1 и v_2, определяемых как:

По нашему опыту, этот вариант критерия с использованием углового преобразования дает менее точные результаты, чем вариант Е.В. Гублера (1978).

$v_1=1;$
$v_2=n_1+n_2-2$

ГЛАВА 6 МЕТОД РАНГОВОЙ КОРРЕЛЯЦИИ

6.1. Обоснование задачи исследования согласованных действий

Первоначальное значение термина "корреляции" - взаимная связь (Oxford Advanced Learner's Dictionary of Current English, 1982). Когда говорят о корреляции, используют термины "корреляционная связь" и "корреляционная зависимость".

Корреляционная связь - это согласованные изменения двух признаков или большего количества признаков (множественная корреляционная связь). Корреляционная связь отражает тот факт, что изменчивость одного признака находится в некотором соответствии с изменчивостью другого (Плохинский Н.А., 1970, с. 40). "Стохастическая[1] связь имеется тогда, когда каждому из значений одной случайной величины соответствует специфическое (условное) распределение вероятностей значений другой величины, и наоборот, каждому из значений этой другой величины соответствует специфическое (условное) распределение вероятностей значений первой случайной величины" (Суходольский Г.В., 1972, с. 178).

Корреляционная зависимость - это изменения, которые вносят значения одного признака в вероятность появления разных значений другого признака.

Оба термина - корреляционная связь и корреляционная зависимость - часто используются как синонимы (Плохинский Н.А.,1970; Суходольский Г.В.,1972; Артемьева Е.Ю., Мартынов Е.М.,1975 и др.). Между тем, согласованные изменения признаков и отражающая это корреляционная связь между ними может свидетельствовать не о зависимости этих признаков между собой, а зависимости обоих этих признаков от какого-то третьего признака или сочетания признаков, не рассматриваемых в исследовании.

Зависимость подразумевает влияние, связь - любые согласованные изменения, которые могут объясняться сотнями причин. Корреляционные связи не могут рассматриваться как свидетельство причинно-следственной связи, они свидетельствуют лишь о том, что изменениям одного признака, как правило, сопутствуют определенные изменения другого, но находится ли причина изменений в одном из признаков или она оказывается за пределами исследуемой пары признаков, нам неизвестно.

Говорить в строгом смысле о зависимости мы можем только в тех случаях, когда сами оказываем какое-то контролируемое воздействие на испытуемых или так организуем исследование, что оказывается возможным точно определить интенсивность не зависящих от нас воздействий. Воздействия, которые мы можем качественно определить или даже измерить, могут рассматриваться как независимые переменные. Признаки, которые мы измеряем и которые, по нашему предположению, могут изменяться под влиянием независимых переменных, считаются зависимыми переменными. Согласованные изменения независимой и зависимой переменной действительно могут рассматриваться как зависимость.

Однако, учитывая, что число градаций, или уровней, зависимой переменной обычно невелико, целесообразнее применять в такого рода исследованиях не корреляционный метод, а методы выявления тенденций изменения признака при изменении условий, например, критерии тенденций H Крускала-Уоллиса и L Пейджа (см. Главы 2 и 3) или метод дисперсионного анализа (см. Главы 7 и 8).

Если в исследование включены независимые переменные, которые мы можем по крайней мере учитывать, например, возраст, то можно считать выявляемые между

---

[1] Стохастическая означает вероятностная. Связи между случайными явлениями называют вероятностными, или стохастическими связями (Суходольский Г. В., 1972, с. 52). Этот термин подчеркивает их отличие от детерминированных или функциональных связей в физике или математике (связь площади треугольника с его высотой и основанием, связь длины окружности с ее радиусом и т. п.). В функциональных связях каждому значению первого признака всегда соответствует (в идеальных условиях) совершенно определенное значение другого признака (Плохинский Н.А., 1970, с. 41). В корреляционных связях каждому значению одного признака может соответствовать определенное распределение значений другого, признака, но не определенное его значение.

возрастом и психологическими признаками корреляционные связи корреляционными зависимостями. В большинстве же случаев нам трудно определить, что в рассматриваемой паре признаков является независимой, а что - зависимой переменной.

Учитывая, что термин "зависимость" явно или неявно подразумевает влияние, лучше пользоваться более нейтральным термином "корреляционная связь".

Корреляционные связи различаются по форме, направлению и степени (силе).

По *форме* корреляционная связь может быть прямолинейной или криволинейной. Прямолинейной может быть, например, связь между количеством тренировок на тренажере и количеством правильно решаемых задач в контрольной сессии. Криволинейной может быть, например, связь между уровнем мотивации и эффективностью выполнения задачи (см. Рис. 6.1). При повышении мотивации эффективность выполнения задачи сначала возрастает, затем достигается оптимальный уровень мотивации, которому соответствует максимальная эффективность выполнения задачи; дальнейшему повышению мотивации сопутствует уже снижение эффективности.



Рис. 6.1. Связь между эффективностью решения задачи и силой мотивационной тенденции (по J.W. Atkinson, 1974, p.200)

По *направлению* корреляционная связь может быть положительной ("прямой") и отрицательной ("обратной"). При положительной прямолинейной корреляции более высоким значениям одного признака соответствуют более высокие значения другого, а более низким значениям одного признака - низкие значения другого (см. Рис. 6.2). При отрицательной корреляции соотношения обратные.

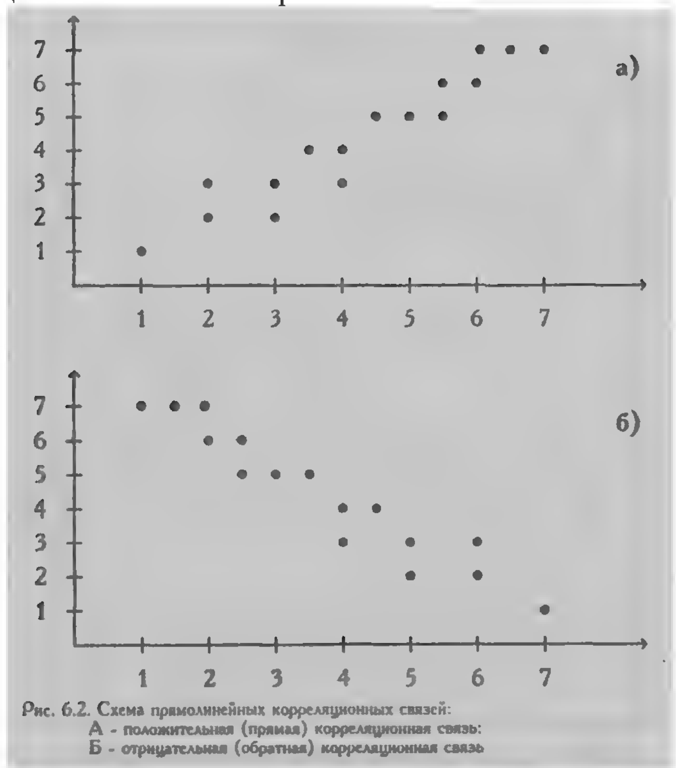


Рис. 6.2. Схема прямолинейных корреляционных связей:
А - положительная (прямая) корреляционная связь:
Б - отрицательная (обратная) корреляционная связь

При положительной корреляции коэффициент корреляции имеет положительный

знак, например *r=+0,207,* при отрицательной корреляции - отрицательный знак, например r=—0,207.

Степень, сила или теснота корреляционной связи определяется по величине коэффициента корреляции.

Сила связи не зависит от ее направленности и определяется по абсолютному значению коэффициента корреляции. Максимальное возможное абсолютное значение коэффициента корреляции r=1,00; минимальное *r=0.*

Используется две системы классификации корреляционных связей по их силе: общая и частная. Общая классификация корреляционных связей (по Ивантер Э.В., Коросову А.В., 1992):

1) сильная, или тесная	при коэффициенте корреляции $r>0,70$;
2) средняя	при $0,50<r<0,69$;
3) умеренная	при $0,30<r<0,49$;
4) слабая	при $0,20<r<0,29$;
5) очень слабая	при $r<0,19$.

Частная классификация корреляционных связей:

1) высокая значимая корреляция	при г, соответствующем уровню статистической значимости $p<0,01$;
2) значимая корреляция	при г, соответствующем уровню статистической значимости $p<0,05$;
3) тенденция достоверной связи	при г, соответствующем уровню статистической значимости $p<0,10$;
4) незначимая корреляция	при г, не достигающем уровня статистической значимости .

Две эти классификации не совпадают. Первая ориентирована только на величину коэффициента корреляции, а вторая определяет, какого уровня значимости достигает данная величина коэффициента корреляции при данном объеме выборки. Чем больше объем выборки, Тем меньшей величины коэффициента корреляции оказьюается достаточно, чтобы корреляция была признана дортоверной. В результате при Малом объеме выборки может оказаться так, что сильная корреляция окажется недостоверной. В то же время при больших объемах выборки Даже слабая корреляция может оказаться достоверной.

Обычно принято ориентироваться на вторую классификацию, поскольку она учитывает объем выборки. Вместе с тем, необходимо помнить, что сильная, или высокая, корреляция - это корреляция с коэффициентом $r>0,70$, а не просто корреляция высокого уровня значимости.

В качестве мер корреляции используются:

1) эмпирические меры тесноты связи, многие из которых были получены еще до открытия метода корреляции, а именно:

а) коэффициент ассоциации, или тетрахорический показатель связи;

б) коэффициенты взаимной сопряженности Пирсона и Чупрова;

в) коэффициент Фехнера;

г) коэффициент корреляции рангов;

2) линейный коэффициент корреляции r,

3) корреляционное отношение η;

4) множественные коэффициенты корреляции и др.

Подробное описание этих мер можно найти в руководствах Ве-нецкого И.Г., Кнльдишева Г.С.(1968), Плохинского Н.А.(1970), Су-ходольского Г.В.(1972), Ивантер Э.В., Коросова А.В.(1992) и др.

В психологических исследованиях чаще всего применяется коэффициент линейной корреляции r Пирсона. Однако этот метод является параметрическим и поэтому не лишен недостатков, свойственных параметрическим методам (см. параграф 1.8).

Параметрическими являются, также методы определения корреляционного отношения и подсчета множественных коэффициентов корреляции. Кроме того, эти методы, как правило, требуют машинной обработки данных. По этим причинам они остаются за пределами нашего рассмотрения.

Все эмпирические меры тесноты связи, кроме коэффициента ранговой корреляции, могут быть заменены методами сопоставления и сравнения, изложенными в Главах 2-5.

Ведь что, в сущности, мы доказываем, когда обосновываем различия в долях двух выборок, характеризующихся исследуемым эффектом? Мы показываем, что если испытуемый относится к одной из выборок, то скорее всего он будет характеризоваться какими-то определенными значениями исследуемого признака, а если он относится к другой из двух выборок, то он будет характеризоваться (с большой степенью вероятности) другими значениями исследуемого признака. Фактически мы исследуем сопряженные изменения двух признаков: *отнесенность* к той или иной выборке и *определенные значения* исследуемого признака.

Что мы доказываем, с другой стороны, когда два распределения признака оказываются сходными или, наоборот, статистически достоверно различающимися между собой? Мы доказываем, что в обеих выборках частоты встречаемости разных значений признака распределяются согласованно или, наоборот, несогласованно.

Мы, правда, скорее определяем меру рассогласованности, чем согласованкости, но все же часто метод χ^2 относится к числу методов, выявляющих степень согласованности или даже связи.

Методы выявления тенденций уже напрямую заменяют меры эмлирической сопряженности, позволяя нам проследить возрастание значений признака при изменении условий. Фактически мы отвечаем на вопрос о том, согласованно ли изменяются условия и значения исследуемого признака.

Быть может, современному психологу не очень просто отказаться от метода подсчета корреляций. Это очень привычно - подсчитывать корреляции. Исторически сложилось так, что этот метод является одним из основных методов статистической обработки. Главное преимущество корреляционного анализа состоит в том, что можно сразу провести множественное сопоставление признаков. Например, "нам необходимо определить, с чем связана успешность в какой-либо деятельности. Исследователь может предполагать, что она связана с уровнем интеллектуального развития, с некоторыми из личностных факторов 16-факторного опросника Кеттелла, а может быть, с уровнем эмпатии, тревожности или фрустрационной толерантности, с возрастом самого испытуемого или возрастом матери в момент его рождения и т.д. и т.п. В итоге он получает связи, отражающие среднегрупповые тенденции сопряженного изменения признаков. Но дело как раз в том, что у каждого отдельного испытуемого успешность в данном виде деятельности может определяться разными психологическими характеристиками или разными их сочетаниями. Метод корреляций отдает предпочтение группе, а не отдельному индивиду.

Против этого можно возразить, что и все остальные статистические методы отдают предпочтение среднегрупповым, а не индивидуальным тенденциям. Однако это не совсем так. Например, метод тенденций L Пейджа определяет степень согласованности *индивидуальных* тенденций, критерий χ^2, Фридмана — степень совпадения или несовпадения *индивидуальных* соотношений рангов, биномиальный критерий m -степень отклонения *индивидуальных* значений от заданных или среднестатистических и т.п.

Прежде чем переходить к корреляциям, исследователю необходимо проанализировать полученные данные с помощью критериев сравнения и сопоставления еще и по другой причине. Возможно, размах вариативности признака в обследованной выборке окажется слишком узким, чтобы можно было распространять полученную корреляцию на весь возможный диапазон его значений. Например, может оказаться так, что в обследованной группе по какому-либо из факторов 16-факторного личностного

опросника Кеттелла получены лишь низкие и средние значения, и в то же время выявлена значимая положительная связь этого личностного фактора с успешностью профессиональной деятельности. Не учитывая истинного размаха значений в данной выборке, можно экстраполировать полученную связь и на высокие значения фактора, что может оказаться ошибкой. Во->первых, связь данного фактора с успешностью деятельности может на самом деле быть криволинейной, как в рассмотренном выше случае связи уровня мотивации с эффективностью выполнения задания (см. Рис. 6.1). Во-вторых, не исключено, что самым важным результатом исследования является как раз факт низких и средних значений данного личностного фактора в обследованной выборке, а исследователь не обратил на него внимания, привычно отдав предпочтение корреляционной матрице, а не таблице первичных данных.

Математическая обработка должна начинаться с использования "самых простых приемов с совершенно понятной для исследователя сутью производимых преобразований" (Дворяшина М.Д., Пехлецкий И.Д., 1976, с. 45). Учитывая большие возможности методов первичной обработки данных, изложенных в Главах 2-5, не исключено, что этими приемами математическая обработка может и заканчиваться. Эти методы дают и основание для достоверных выводов, и материал для выдвижения новых гипотез, и стимул к новым размышлениям.

И все же, если исследователь хочет применить метод корреляций, в настоящем пособии предлагается использовать коэффициент ранговой корреляции Спирмена. Основанием для выбора этого коэффициента служат:

а) его универсальность;

б) простота;

в) широкие возможности в решении задач сравнения индивидуальных или групповых иерархий признаков.

Универсальность коэффициента ранговой корреляции проявляется в том, что он применим к любым количественно измеренным или ранжированным данным. Простота метода позволяет подсчитывать корреляцию "вручную". Уникальность метода ранговой корреляции состоит в том, что он позволяет сопоставлять не индивидуальные показатели, а индивидуальные иерархии, или профили, что недоступно ни одному из других статистических методов, включая метод линейной корреляции (Плохинский НА., 1970, с. 167).

Коэффициент ранговой корреляции рекомендуется применять в тех случаях, когда нам необходимо проверить, согласованно ли изменяются разные признаки у одного и того же испытуемого и насколько совпадают индивидуальные ранговые показатели у двух отдельных испытуемых или у испытуемого и группы.

6.2. Коэффициент ранговой корреляции r_s Спирмена

Назначение рангового коэффициента корреляции

Метод ранговой корреляции Спирмена позволяет определить тесноту (силу) и направление корреляционной связи между *двумя признаками* или *двумя профилями {иерархиями)* признаков.

Описание метода

Для подсчета ранговой корреляции необходимо располагать двумя рядами значений, которые могут быть проранжированы. Такими рядами значений могут быть:

1) *два признака,* измеренные в одной и той же группе испытуемых;

2) *две индивидуальные иерархии признаков,* выявленные у двух испытуемых по одному и тому же набору признаков (например, личностные профили по 16-факторному опроснику Р. Б. Кеттелла, иерархии ценностей по методике Р. Рокича, последовательности предпочтений в выборе из нескольких альтернатив и др.);

3) *две групповые иерархии признаков;*

4) *индивидуальная и групповая* иерархии признаков.

Вначале показатели ранжируются отдельно по каждому из признаков. Как правило, меньшему значению признака начисляется меньший ранг.

Рассмотрим случай 1 (два признака). Здесь ранжируются индивидуальные значения по первому признаку, полученные разными испытуемыми, а затем индивидуальные значения по второму признаку.

Если два признака связаны положительно, то испытуемые, имеющие низкие ранги по одному из них, будут иметь низкие ранги и по другому, а испытуемые, имеющие высокие ранги по одному из признаков, будут иметь по другому признаку также высокие ранги. Для подсчета r_s необходимо определить разности (d) между рангами, полученными данным испытуемым по обоим признакам. Затем эти показатели d определенным образом преобразуются и вычитаются из 1. Чем меньше разности между рангами, тем больше будет r_s, тем ближе он будет к +1.

Если корреляция отсутствует, то все ранги будут перемешаны и между ними не будет никакого соответствия. Формула составлена так, что в этом случае r_s, окажется близким к 0.

В случае отрицательной корреляции низким рангам испытуемых по одному признаку будут соответствовать высокие ранги по другому признаку, и наоборот.

Чем больше несовпадение между рангами испытуемых по двумя переменным, тем ближе r_s к -1.

Рассмотрим случай 2 (два индивидуальных профиля). Здесь ранжируются индивидуальные значения, полученные каждым из 2-х испытуемым по определенному (одинаковому для них обоих) набору признаков. Первый ранг получит признак с самым низким значением; второй ранг - признак с более высоким значением и т.д. Очевидно, что все признаки должны быть измерены в одних и тех же единицах, иначе ранжирование невозможно. Например, невозможно проранжировать показатели по личностному опроснику Кеттелла (16*PF*), если они выражены в "сырых" баллах, поскольку по разным факторам диапазоны значений различны: от 0 до 13, от 0 до 20 и от 0 до 26. Мы не можем сказать, какой из факторов будет занимать первое место по выраженности, пока не приведем все значения к единой шкале (чаще всего это шкала стенов).

Если индивидуальные иерархии двух испытуемых связаны положительно, то признаки, имеющие низкие ранги у одного из них, будут иметь низкие ранги и у другого, и наоборот. Например, если у одного испытуемого фактор Е (доминантность) имеет самый низкий ранг, то и у другого испытуемого он должен иметь низкий ранг, если у одного испытуемого фактор С (эмоциональная устойчивость) имеет высший ранг, то и другой испытуемый должен иметь по этому фактору высокий ранг и т.д.

Рассмотрим случай 3 (два групповых профиля). Здесь ранжируются среднегрупповые значения, полученные в 2-х группах испытуемых по определенному, одинаковому для двух групп, набору признаков. В дальнейшем линия рассуждений такая же, как и в предыдущих двух случаях.

Рассмотрим случай 4 (индивидуальный и групповой профили). Здесь ранжируются отдельно индивидуальные значения испытуемого и среднегрупповые значения по тому же набору признаков, которые получены, как правило, при исключении этого отдельного испытуемого - он не участвует в среднегрупповом профиле, с которым будет сопоставляться его индивидуальный профиль. Ранговая корреляция позволит проверить, насколько согласованы индивидуальный и групповой профили.

Во всех четырех случаях значимость полученного коэффициента корреляции определяется по количеству ранжированных значений *N*. В первом случае это количество будет совпадать с объемом выборки n. Во втором случае количеством наблюдений будет количество признаков, составляющих иерархию. В третьем и четвертом случае *N* - это

также количество сопоставляемых признаков, а не количество испытуемых в группах. Подробные пояснения даны в примерах.

Если абсолютная величина r_s достигает критического значения или превышает его, корреляция достоверна.

Гипотезы

Возможны два варианта гипотез. Первый относится к случаю 1, второй - к трем остальным случаям.

Первый вариант гипотез

H_0: Корреляция между переменными А и Б не отличается от нуля.

H_1: Корреляция между переменными А и Б достоверно отличается от нуля.

Второй вариант гипотез

H_0: Корреляция между иерархиями А и Б не отличается от нуля.

H_1: Корреляция между иерархиями А и Б достоверно отличается от нуля.

Графическое представление метода ранговой корреляции

Чаще всего корреляционную связь представляют графически в виде облака точек или в виде линий, отражающих общую тенденцию размещения точек в пространстве двух осей: оси признака А и признака Б (см. Рис. 6.2).

Попробуем изобразить ранговую корреляцию в виде двух рядов ранжированных значений, которые попарно соединены линиями (Рис. 6.3). Если ранги по признаку А и по признаку Б совпадают, то между ними оказывается горизонтальная линия, если ранги не совпадают, то линия становится наклонной. Чем больше несовпадение рангов, тем более наклонной становится линия. Слева на Рис. 6.3 отображена максимально высокая положительная корреляция ($r_в$=+1,0) - практически это "лестница". В центре отображена нулевая корреляция - плетенка с неправильными переплетениями. Все ранги здесь перепутаны. Справа отображена максимально высокая отрицательная корреляция (r_s=-1,0) -паутина с правильным переплетением линий.

Рис. 6.3. Графическое представление ранговой корреляции:

а) высокая положительная корреляция;

б) нулевая корреляция;

в) высокая отрицательная корреляция

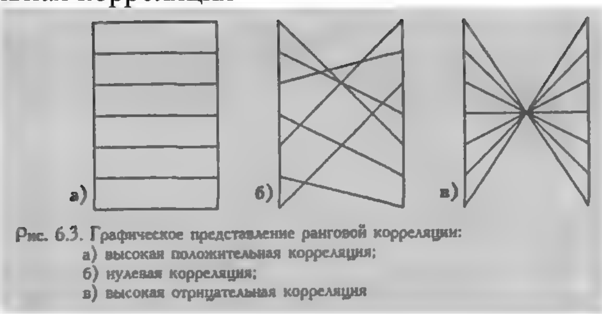


Рис. 6.3. Графическое представление ранговой корреляции:
а) высокая положительная корреляция;
б) нулевая корреляция;
в) высокая отрицательная корреляция

Ограничения коэффициента ранговой корреляции

1. По каждой переменной должно быть представлено не менее 5 наблюдений. Верхняя граница выборки определяется имеющимися таблицами критических значений (Табл.XVI Приложения 1), а именно $N \leq 40$.

2. Коэффициент ранговой корреляции Спирмена r_s при большом количестве одинаковых рангов по одной или обеим сопоставляемым переменным дает огрубленные значения. В идеале оба коррелируемых ряда должны представлять собой две последовательности несовпадающих значений. В случае, если это условие не соблюдается, необходимо вносить поправку на одинаковые ранги. Соответствующая формула дана в примере 4.

Пример 1 - корреляция между двумя признаками

В исследовании, моделирующем деятельность авиадиспетчера (Одерышев Б.С., Шамова Е.П., Сидоренко Е.В., Ларченко Н.Н., 1978), группа испытуемых, студентов физического факультета ЛГУ проходила подготовку перед началом работы на тренажере. Испытуемые должны были решать задачи по выбору оптимального типа взлетно-посадочной полосы для заданного типа самолета. Связано ли количество ошибок, допущенных испытуемыми в тренировочной сессии, с показателями вербального и невербального интеллекта, измеренными по методике Д. Векслера?

Таблица 6.1

Показатели количества ошибок в тренировочной сессии и показатели уровня вербального и невербального интеллекта у студентов-физиков (N=10)

Испытуемый		Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	Т.А.	29	131	106
2	П.А.	54	132	90
3	Ч.И.	13	121	95
4	Ц.А.	8	127	116
5	См.А.	14	136	. 127
6	К.Е.	26	124	107
7	К.А.	9	134	104
8	Б.Л.	20	136	102
9	И.А.	2	132	111
10	Ф.В.	17	136	99
Суммы		192	1309	1057
Средние		19,2	130,9	105,7

Сначала попробуем ответить на вопрос, связаны ли между собой показатели количества ошибок и вербального интеллекта.

Сформулируем гипотезы.

H_0: Корреляция между показателем количества ошибок в тренировочной сессии и уровнем вербального интеллекта не отличается от нуля.

H_1: Корреляция между показателем количества ошибок в тренировочной сессии и уровнем вербального интеллекта статистически значимо отличается от нуля.

Далее нам необходимо проранжировать оба показателя, Приписывая меньшему значению меньший ранг, затем подсчитать разности между рангами, которые получил каждый испытуемый по двум переменным (признакам), и возвести эти разности в квадрат. Произведем все необходимые расчеты в таблице.

В Табл. 6.2 в первой колонке слева представлены значения по показателю количества ошибок; в следующей колонке - их ранги. В третьей колонке слева представлены значения по показателю вербального интеллекта; в следующем столбце - их ранги. В пятом слева представлены разности d между рангом по переменной А (количество ошибок) и переменной Б (вербальный интеллект). В последнем столбце представлены квадраты разностей - d^2.

Таблица 6.2

Расчет d^2 для рангового коэффициента корреляции Спирмена r_s при сопоставлении показателей количества ошибок и вербального интеллекта у студентов-физиков (N=10)

Испытуемый		Переменная А количество ошибок		Переменная Б вербальный интеллект.		*d* (ранг А - - ранг Б)	J^2
		Индивидуальные значения	Ранг	Индивидуальные значения	Ранг		
1	ТА.	29	9	131	4	5	25
2	ПА.	54	10	132	5.5	4,5	20.25
3	Ч.И.	13	4	121	1	3	9
4	Ц.А.	8	2	127	3	-1	1
5	См.А.	14	5	136	9	-4	16
6	К.Е.	26	8	124	2	6	36

7	К.А.	9	3	134	7	-4	16
8	Б.Л.	20	7	136	9	-2	4
9	И.А.	2	1	132	5,5	-4,5	20,25
10	Ф.В.	17	6	136	9		9
Суммы			55		55	0	156,5

Коэффициент ранговой корреляции Спирмена подсчитывается по формуле:

$$r_s = 1 - \frac{6 \cdot \sum(d^2)}{N \cdot (N^2 - 1)}$$

где d - разность между рангами по двум переменным для каждого испытуемого;

N - количество ранжируемых значений, в. данном случае количество испытуемых.

Рассчитаем эмпирическое значение r_s:

$$r_s = 1 - \frac{6 \cdot 156{,}5}{10 \cdot (10^2 - 1)} = 1 - \frac{939}{990} = 0{,}052$$

Полученное эмпирическое значение r_s близко к 0. И все же определим критические значения r_s при N=10 по Табл. XVI Приложения 1:

$$r_{s\ кр.} = \begin{cases} 0{,}64\ (\rho \le 0.05) \\ 0{,}79\ (\rho \le 0.01) \end{cases}$$

$r_{s\ эмп} < r_{s\ кр.}$

Ответ: H_0 принимается. Корреляция между показателем количества ошибок в тренировочной сессии и уровнем вербального интеллекта не отличается от нуля.

Теперь попробуем ответить на вопрос, связаны ли между собой показатели количества ошибок и невербального интеллекта.

Сформулируем гипотезы.

H_0: Корреляция между показателем количества ошибок в тренировочной сессии и уровнем невербального интеллекта не отличается от 0.

H_1: Корреляция между показателем количества ошибок в тренировочной сессии и уровнем невербального интеллекта статистически значимо отличается от 0.

Результаты ранжирования и сопоставления рангов представлены в Табл. 6.3.

Таблица 6.3

Расчет d^2 для рангового коэффициента корреляции Спирмена r_s при сопоставлении показателей количества ошибок и невербального интеллекта у студентов-физиков (N=10)

Испытуемый		Переменная А количество ошибок		Переменная Е ; невербальный интеллект		*d* (ранг А — — ранг Б)	d^2
		Индивидуальные значения	Ранг	Индивидуальные значения	Ранг		
1	Т.А.	29	9	106	6	3	9
2	П:А.	54	10	90	1	9	81
3	Ч.И.	13	4	95	2	2	4
4	Ц-А. „	8	2	116	9	-7	49
5	См.А.	14	5	127	10	-5	25
6	К.Е.	26	8	107	7	1	1
7	К.А.	9	3	104	5	-2	4
8	Б.Л.	20	7	102	4	3	9
9	И.А.	2	1	111	8	-7	49
10	Ф.В.	17	6	99	3	3	9
Суммы			55		55	0	240

Рассчитаем эмпирическое значение r_s:

$$r_s = 1 - \frac{6 \cdot 240}{10 \cdot (10^2 - 1)} = 1 - \frac{1440}{990} = -0{,}455$$

Критические значения те же, что и в предыдущем случае:

$$r_{s\ кр.} = \begin{cases} 0{,}64\ (\rho \le 0.05) \\ 0{,}79\ (\rho \le 0.01) \end{cases}$$

Мы помним, что для определения значимости r_s неважно, является ли он положительным или отрицательным, важна лишь его абсолютная величина. В данном случае:

$r_{s\ эмп} < r_{s\ кр}$.

Ответ: H_0 принимается. Корреляция между показателем количества ошибок в тренировочной сессии и уровнем невербального интеллекта случайна, r_s не отличается от 0.

Вместе с тем, мы можем обратить внимание на определенную тенденцию *отрицательной* связи между этими двумя переменными. Возможно, мы смогли бы ее подтвердить на статистически значимом уровне, если бы увеличили объем выборки.

Пример 2 - корреляция между индивидуальными профилями

В исследовании, посвященном проблемам ценностной реориента-ции, выявлялись иерархии терминальных ценностей по методике М. Рокича у родителей и их взрослых детей (Сидоренко Е.В., 1996). Ранги терминальных ценностей, полученные при обследовании пары мать-дочь (матери - 66 лет, дочери - 42 года) представлены в Табл. 6.4. Попытаемся определить, как эти ценностные иерархии коррелируют друг с другом.

Таблица 6.4

Ранги терминальных ценностей по списку М.Рокича в индивидуальных иерархиях матери и дочери

Терминальные ценности	Ряд1: Ранг ценностей в иерархии матери	Ряд 2: Ранг ценностей в иерархии дочери	d	d^2
1 Активная деятельная жизнь	15	15	0	0
2 Жизненная мудрость	1	3	-2	4
3 Здоровье	7	14	-7	49
4 Интересная работа	8	12	-4	16
5 Красота природы и искусство	16	17	-1	1
6 Любовь	11	10	1	1
7 Материально обеспеченная жизнь	12	13	-1	1
8 Наличие хороших и верных друзей	9	11	-2	4
9 Общественное признание	17	5	12	144
10 Познание	5	1	4	16
11 Продуктивная жнзнь	2	2	0	0
12 Развитие	6	8	-2	4
13 Развлечения	18	18	0	0
14 Свобода	4	6	-2	4
15 Счастливая семейная жизнь	13	4	9	81
16 Счастье других	14	16	-2	4
17 Творчество	10	9	1	1
18 Уверенность в себе	3	7	-4	16
Суммы	171	171	0	346

Сформулируем гипотезы.

H_0: Корреляция между иерархиями терминальных ценностей матери и дочери не отличается от нуля.

H_1: Корреляция между иерархиями терминальных ценностей матери и дочери статистически значимо отличается от нуля.

Поскольку ранжирование ценностей предполагается самой процедурой исследования, нам остается лишь подсчитать разности между рангами 18 ценностей в двух иерархиях . В 3-м и 4-м столбцах Табл. 6.4 представлены разности d и квадраты этих разностей d^2.

Определяем эмпирическое значение r_s по формуле:

---

[2] Обычно рекомендуется всегда меньшему значению приписывать меньший ранг (см. Пример 1). В данном случае самая значимая ценность получает меньший ранг. Для подсчета коэффициента это несущественно. Главное, чтобы ранжирование было в обоих рядах *однонаправленным.*

$$r_s = 1 - \frac{6 \cdot \Sigma(d^2)}{N \cdot (N^2 - 1)}$$

где d - разности между рангами по каждой из переменных, в данном случае по каждой из терминальных ценностей;

N - количество переменных, образующих иерархию, в данном случае количество ценностей.

Для данного примера:

$$r_s = 1 - \frac{6 \cdot 346}{18 \cdot (18^2 - 1)} = 1 - \frac{2076}{5814} = 0.643$$

По Табл. XVI Приложения 1 определяем критические значения:

$$r_{s\ кр.} = \begin{cases} 0{,}47\ (\rho \le 0.05) \\ 0.60\ (\rho \le 0.01) \end{cases}$$

$$r_{s\ эмп} > r_{s\ кр}\ (\rho \le 0{,}01)$$

Ответ: H_0 отвергается. Принимается H_1. Корреляция между иерархиями терминальных ценностей матери и дочери статистически значима ($p<0{,}01$) и является положительной.

По данным Табл. 6.4 мы можем определить, что основные расхождения приходятся на ценности "Счастливая семейная жизнь", "Общественное признание" и "Здоровье", ранги остальных ценностей достаточно близки.

Пример 3 - корреляция между двумя групповыми иерархиями

Джозеф Вольпе в книге, написанной совместно с сыном (Wolpe J., Wolpe D., 1981) приводит упорядоченный перечень из наиболее часто встречающихся у современного человека "бесполезных", по его обозначению, страхов, которые не несут сигнального значения и лишь мешают полноценно жить и действовать. В отечественном исследовании, проведенном М.Э. Раховой (1994) 32 испытуемых должны были по 10-балльной шкале оценить, насколько актуальным для них является тот или иной вид страха из перечня Вольпе[3]. Обследованная выборка состояла из студентов Гидрометеорологического и Педагогического институтов Санкт-Петербурга: 15 юношей и 17 девушек в возрасте от 17 до 28 лет, средний возраст 23 года.

Данные, полученные по 10-балльной шкале, были усреднены по 32 испытуемым, и средние проранжированы. В Табл. 6.5 представлены ранговые показатели, полученные Дж. Вольпе и М. Э. Раховой. Совпадают ли ранговые последовательности 20 видов страха?

Сформулируем гипотезы.

H_0: Корреляция между упорядоченными перечнями видов страха в американской и отечественных выборках не отличается от нуля.

H_1: Корреляция между упорядоченными перечнями видов страха в американской и отечественной выборках статистически значимо отличается от нуля.

Все расчеты, связанные с вычислением и возведением в квадрат разностей между рангами разных видов страха в двух выборках, представлены в Табл. 6.5.

Таблица 6.5

Расчет d для рангового коэффициента корреляции Спирмена при сопоставлении упорядоченных перечней видов страха в американской и отечественной выборках

	Виды страха	Ранг в американской выборке	Ранг в российской выборке	d	d^2
1	Страх публичного выступления	1	7	-6	36
2	Страх полета	2	12	-10	100

[3] В исследовании М.Э. Раховой были выявлены виды страха, отсутствующие в перечне Вольпе, например, страх за благополучие близких (1-й ранг), неизвестности (5-й ранг), нападения (8-й ранг) и др. Однако в данном примере в ранжировании участвуют только 20 страхов из перечня Вольпе, поскольку мы можем подсчитывать коэффициент корреляции лишь между теми признаками, которые измерены в обеих выборках.

3	Страх совершить ошибку	3	10	-7	49
4	Страх неудачи	4	6	-2	4
5	Страх неодобрения	5	9	-4	16
6	Страх отвержения	6	2	4	16
7	Страх злых люден	7	5	2	4
8	Страх одиночества	8	1	7	49
9	Страх крови	9	16	-7	49
10	Страх открытых ран	10	13	-3	9
И	Страх дантиста	11	3	8	64
12	Страх уколов	12	19	-7	49
13	Страх прохождения тестов	13	20	-7	49
14	Страх полиции ^милиции)	14	17	-3	9
15	Страх высоты	15	4	11	121
16	Страх собак	16	11	5	25
17	Страх пауков	17	18	-1	1
18	Страх искалеченных людей	18	8	10	100
19	Страх больниц	19	15	4	16
20	Страх темноты	20	14	6	36
	Суммы	210	210	0	802

Определяем эмпирическое значение r_s:

$$r_s = 1 - \frac{6 \cdot \Sigma(d^2)}{N \cdot (N^2 - 1)} = 1 - \frac{6 \cdot 802}{20 \cdot (20^2 - 1)} = 1 - \frac{4812}{7980} = 0{,}397$$

По Табл. XVI Приложения 1 определяем критические значения r_s при N=20:

$$r_{s\ кр.} = \begin{cases} 0{,}45\ (\rho \leq 0{,}05) \\ 0{,}57\ (\rho \leq 0{,}01) \end{cases}$$

$r_{s\ эмп} < r_{s\ кр.}$

Ответ: H_0 принимается. Корреляция между упорядоченными перечнями видов страха в американской и отечественной выборках не достигает уровня статистической значимости, т. е. значимо не отличается от нуля.

Пример 4 - корреляция между индивидуальным и среднегрупповым профилями

Выборке петербуржцев в возрасте от 20 до 78 лет (31 мужчина, 46 женщин), уравновешенной по возрасту таким образом, что лица в возрасте старше 55 лет составляли в ней 50%[4], предлагалось ответить на вопрос: "Какой уровень развития каждого из перечисленных ниже качеств необходим для депутата Городского собрания Санкт-Петербурга?" (Сидоренко Е.В., Дерманова И.Б., Анисимова О.М., Витенберг Е.В., Шульга А.П., 1994). Оценка производилась по 10-балльной шкале. Параллельно с этим обследовалась выборка из депутатов и кандидатов в депутаты в Городское собрание Санкт-Петербурга (n=14). Индивидуальная диагностика политических деятелей и претендентов производилась с помощью Оксфордской системы экспресс-видеодиагностики по тому же набору личностных качеств, который предъявлялся выборке избирателей.

В Табл. 6.6 представлены средние значения, полученные для каждого из качеств **в** выборке избирателей ("эталонный ряд") и индивидуальные значения одного из депутатов Городского собрания.

Попытаемся определить, насколько индивидуальный профиль депутата К-ва коррелирует с эталонным профилем.

Таблица 6.6

Усредненные эталонные оценки избирателей (п=77) и индивидуальные показатели депутата К-ва по 18 личностным качествам экспресс-видеодиагностики

Наименование качества	Усредненные эталонные оценки избирателей	Индивидуальные показатели депутата К-ва

[4] Введение этого условия диктовалось тем, что в непосредственно предшествовавших исследованию выборах 52% электората составляли лица старше 55 лет.

1. Общий уровень культуры	8,64	15
2. Обучаемость	7,89	7
3. Логика	8,38	12
4. Способность к творчеству нового	6,97	5
5.. Самокритичность	8,28	14
6. Ответственность	9,56	18
7. Самостоятельность	8,12	13
8. Энергия, активность	8,41	17
9. Целеустремленность	8,00	19
10. Выдержка, самообладание	8,71	9
И. Стойкость	7,74	16
12. Личностная зрелость	8,10	11
13. Порядочность	9,02	12
14. Гуманизм	7.89	10
15. Умение общаться с людьми	8,74	8
16. Терпимость к чужому мнению	7.84	6
17. Гибкость поведения	7,67	4
18. Способность производить благоприятное впечатление	7,23	8

Таблица 6.7

Расчет d^2 для рангового коэффициента корреляции Спирмена между эталонным и индивидуальным профилями личностных качеств депутата

Наименование качества	Ряд 1: ранг качества в эталонном профиле	Ряд 2: ранг качества в индивидуальном профиле	d	d^2
1 Ответственность	1	2	-1	1
2 Порядочность	2	8,5	-6,5	42,25
3 Умение общаться с людьми	3	13,5	-10,5	110,25
4 Выдержка, самообладание	4	12	-8	64
5 Общий уровень культуры	5	5	0	0
6 Энергия, активность	6	3	3	9
7 Логика	7	8,5	-1,5	2,25
8 Самокритичность	8	6	2	4
9 Самостоятельность	9	7	2	4
10 Личностная зрелость	10	10	0	0
И Целеустремленность	И	1	10	100
12 Обучаемость	12,5	15	-2,5	6,25
13 Гуманизм	12,5	И	1,5	2,25
14 Терпимость к чужому мнению	14	16	-2	4
15 Стойкость	15	4	11	121
16 Гибкость поведения	16	18	-2	4
17 Способность производить благоприятное впечатление	17	13,5	3,5	12,25
18 Способность к творчеству нового	18.	17	1	1
Суммы	171	171	0	487,5

Как видно из Табл. 6.6, оценки избирателей и индивидуальные показатели депутата варьируют в разных диапазонах. Действительно оценки избирателей были получены по 10-балльной шкале, а индивидуальные показатели по экспресс-видеодиагностике измеряются по 20-ти балльной шкале. Ранжирование позволяет нам перевести обе шкалы измерения в единую шкалу, где единицей измерения будет 1 ранг, а максимальное значение составит 18 рангов.

Ранжирование, как мы помним, необходимо произвести отдельно по каждому ряду значений. В данном случае целесообразно начислять большему значению меньший ранг, чтобы сразу можно было увидеть, на каком месте по значимости (для избирателей) или по выраженности (у депутата) находится то или иное качество.

Результаты ранжирования представлены в Табл. 6.7. Качества перечислены в последовательности, отражающей эталонный профиль.

Сформулируем гипотезы.

H_0: Корреляция между индивидуальным профилем депутата К-ва и эталонным профилем, построенным по оценкам избирателей, не отличается от нуля.

H_1: Корреляция между индивидуальным профилем депутата К-ва и эталонным профилем, построенным по оценкам избирателей, статистически значимо отличается от нуля. Поскольку в обоих сопоставляемых ранговых рядах присутствуют

группы одинаковых рангов, перед подсчетом коэффициента ранговой

корреляции необходимо внести поправки на одинаковые ранги T_a и T_b:

$$T_a=\Sigma(a^3-a)/12$$
$$T_b=\Sigma(b^3-b)/12$$

где a - объем каждой группы одинаковых рангов в ранговом ряду А,

b - объем каждой группы одинаковых рангов в ранговом ряду В.

В данном случае, в ряду А (эталонный профиль) присутствует одна группа одинаковых рангов - качества "обучаемость" и "гуманизм" имеют один и тот же ранг 12,5; следовательно, a=2.

$T_a=(2^3-2)/12=0,50$.

В ряду В (индивидуальный профиль) присутствует две группы одинаковых рангов, при этом $b_1=2$ и $b_2=2$.

$T_a=[(2^3-2)+(2^3-2)]/12=1,00$

Для подсчета эмпирического значения r_s используем формулу

$$r_s = 1-6\cdot\frac{\sum d^2+T_a+T_b}{N\cdot(N^2-1)}$$

В данном случае:

$$r_s = 1-\frac{6\cdot 487,5+0,50+1,00}{18\cdot(18^2-1)} = 1-\frac{2926,5}{5814} = 0,4967$$

Заметим, что если бы поправка на одинаковые ранги нами не вносилась, то величина r_s была бы лишь на (на 0,0002) выше:

$$r_s = 1-\frac{6\cdot 487,5}{18\cdot(18^2-1)} = 1-\frac{2925}{5814} = 0,4969$$

При больших количествах одинаковых рангов изменения r_5 могут оказаться гораздо более существенными. Наличие одинаковых рангов означает меньшую степень дифференцированное™ упорядоченных переменных и, следовательно, меньшую возможность оценить степень связи между ними (Суходольский Г.В., 1972, с.76).

По Табл. XVI Приложения 1 определяем критические значения г, при N=18:

$$r_{s\ кр.} = \begin{cases} 0,47\ (\rho\le 0,05) \\ 0.60\ (\rho\le 0,01)\end{cases}$$

$$r_{s\ эмп} > r_{s\ кр.}\ (\rho\le 0,05)$$

Ответ: HQ отвергается. Корреляция между индивидуальным профилем депутата К-ва и эталонным профилем, отвечающим требованиям избирателей, статистически значима ($p<0,05$) и является положительной.

Из Табл. 6.7 видно, что депутат К-в имеет более низкий ранг по шкалам Умения общаться с людьми и более высокие ранги по шкалам Целеустремленности и Стойкости, чем это предписывается избирательским эталоном. Этими расхождениями, главным образом, и объясняется некоторое снижение полученного r_s.

Сформулируем общий алгоритм подсчета r_s.

АЛГОРИТМ 20

Расчет коэффициента ранговой корреляции Спирмена r_s.

1. Определить, какие два признака или две иерархии признаков будут участвовать в сопоставлении как переменные А и В.

2. Проранжировать значения переменной А, начисляя ранг 1 наименьшему

значению, в соответствии с правилами ранжирования (см. п.2.3). Занести ранги в первый столбец таблицы по порядку номеров испытуемых или признаков.

3. Проранжировать значения переменной В, в соответствии с теми же правилами. Занести ранги во второй столбец таблицы по порядку номеров испытуемых или признаков.

4. Подсчитать разности *d* между рангами А и В по каждой строке таблицы и занести в третий столбец таблицы.

5. Возвести каждую разность в квадрат: d^2 . Эти значения занести в четвертый столбец таблицы.

6. Подсчитать сумму квадратов $\sum d^2$.

7. При наличии одинаковых рангов рассчитать поправки:

$T_a = \Sigma(a^3 - a)/12$

$T_b = \Sigma(b^3 - b)/12$

где *a* - объем каждой группы одинаковых рангов в ранговом ряду А;

b - объем каждой группы одинаковых рангов в ранговом ряду В.

8. Рассчитать коэффициент ранговой корреляции r_5 по формуле:

а) при отсутствии одинаковых рангов

$$r_s = 1 - 6 \cdot \frac{\sum d^2}{N \cdot (N^2 - 1)}$$

б) при наличии одинаковых рангов

$$r_s = 1 - 6 \cdot \frac{\sum d^2 + T_a + T_b}{N \cdot (N^2 - 1)}$$

где $\sum d^2$ - сумма квадратов разностей между рангами*

T_a и T_b, - поправки на одинаковые ранги;

N - количество испытуемых или признаков, участвовавших в ранжировании.

9. Определить по Табл. XVI Приложения 1 критические значения r_s для данного N. Если r_s превышает критическое значение или по крайней мере равен ему, корреляция достоверно отличается от 0.

ГЛАВА 7 ДИСПЕРСИОННЫЙ АНАЛИЗ

7.1. Понятие дисперсионного анализа

Дисперсионный анализ - это анализ изменчивости признака под влиянием каких-либо контролируемых переменных факторов. В зарубежной литературе дисперсионный анализ часто обозначается как ANOVA, что переводится как анализ вариативности (Analysis of Variance). Автором метода является Р. А. Фишер (Fisher R.A., 1918, 1938).

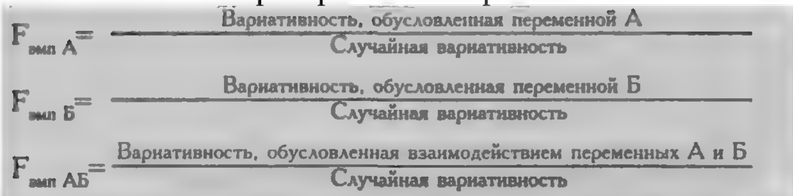
Задача дисперсионного анализа состоит в том, чтобы из общей вариативности признака вычленить вариативность троякого рода:

а) вариативность, обусловленную *действием каждой* из исследуемых независимых переменных;

б) вариативность, обусловленную *взаимодействием* исследуемых независимых переменных;

в) *случайную* вариативность, обусловленную всеми другими неизвестными переменными.

Вариативность, обусловленная действием исследуемых переменных и их взаимодействием, соотносится со случайной вариативностью. Показателем этого соотношения является критерий F Фишера[1].



В формулу расчета критерия F входят оценки дисперсий, то есть параметров распределения признака, поэтому критерий F является параметрическим критерием.

---

[1] Критерии F Фишера и метод углового преобразования Фишера, дающий нам критерий φ*, - это совершенно различные методы, имеющие разное предназначение и разные способы вычисления.

Чем в большей степени вариативность признака обусловлена исследуемыми переменными (факторами) или их взаимодействием, тем выше эмпирические значения критерия F.

В дисперсионном анализе исследователь исходит из предположения, что одни переменные могут рассматриваться как причины, а другие - как следствия. Переменные первого рода считаются факторами, а переменные второго рода - результативными признаками. В этом отличие дисперсионного анализа от прямолинейного корреляционного анализа, в котором мы исходим из предположения, что изменения одного признака просто сопровождаются определенными изменениями другого.

В дисперсионном анализе возможны два принципиальных пути разделения всех исследуемых переменных на независимые переменные (факторы) и зависимые переменные (результативные признаки).

Первый путь состоит в том, что мы совершаем какие-либо воздействия на испытуемых или учитываем какие-либо не зависящие от нас воздействия на них, и именно эти воздействия считаем независимыми переменными, или факторами, а исследуемые признаки рассматриваем как зависимые переменные, или результативные признаки. Например, возраст испытуемых или способ предъявления им информации считаем факторами, а обучаемость или эффективность выполнения задания - результативными признаками.

Второй путь предполагает, что мы, не совершая никаких воздействий, считаем, что при разных уровнях развития одних психологических признаков другие проявляются тоже по-разному. По тем или иным причинам мы решаем, что одни признаки могут рассматриваться скорее как факторы, а другие - как результат действия этих факторов.

Например, уровень интеллекта или мотивации достижения начинаем считать факторами, а профессиональную компетентность или социометрический статус - результативными признаками.

Второй путь весьма уязвим для критики. Допустим, мы предположили, что настойчивость - значимый фактор учебной успешности студентов. Мы принимаем настойчивость за воздействующую переменную (фактор), а учебную успешность - за результативный признак. Против этого могут быть выдвинуты сразу же два возражения. Во-первых, успех может стимулировать настойчивость; во-вторых, как, собственно, измерялась настойчивость? Если она измерялась с помощью метода экспертных оценок, а экспертами были соученики или преподаватели, которым известна учебная успешность испытуемых, то не исключено, что это оценка настойчивости будет зависеть от известных экспертам показателей успешности, а не наоборот.

Допустим, что в другом исследовании мы исходим из предположения, что фактор социальной смелости (фактор Н) из 16-факторного личностного опросника Р.Б. Кеттелла - это та независимая переменная, которая определяет объем заключенных торговым представителем договоров на поставку косметических товаров. Но если объем договоров определялся по какому-то периоду работы, скажем трехмесячному, а личностное обследование проводилось в конце этого периода или даже после его истечения, то мы не можем со всей уверенностью отделить здесь причину от следствия. Есть очень сильное направление в психологии и психотерапии, которое утверждает, что личностные изменения начинаются с действий и поступков: "Начни действовать, и постепенно станешь таким, как твои поступки". Таким образом, психолог, представляющий это направление, возможно, стал бы утверждать, что причиной должен считаться достигнутый объем договорных поставок, а результатом - повышение социальной смелости.

Только наше исследовательское чутье может подсказать нам, что должно рассматриваться как причина, а что - как результат. Однако не всегда эти ощущения у разных исследователей совпадают, поэтому нужно быть готовым к тому, что наши выводы могут быть оспорены другими специалистами, которые рассматривают данный предмет с иной точки зрения и видят в нем иные перспективы. Впрочем, спорность выводов - постоянный спутник психологического исследования.

Постараемся быть оптимистичными и представим себе, что существует все же какое-то совпадение взглядов на психологические причины и следствия. На Рис. 7.1 представлены два варианта рассеивания показателей учебной успешности в зависимости от уровня развития кратковременной памяти. Из Рис. 7.1(а) мы видим, что при низком уровне развития кратковременной памяти оценки по английскому языку, похоже, несколько ниже, чем при среднем, а при высоком уровне выше, чем при среднем. Похоже, что кратковременная память может рассматриваться как фактор успешности овладения английским языком. С другой стороны, Рис. 7.1(6) свидетельствует о том, что успешность в чистописании вряд ли так же определенно зависит от уровня развития кратковременной памяти.

О том, верны ли наши предположения, мы сможем судить только после вычисления эмпирических значений критерия F.

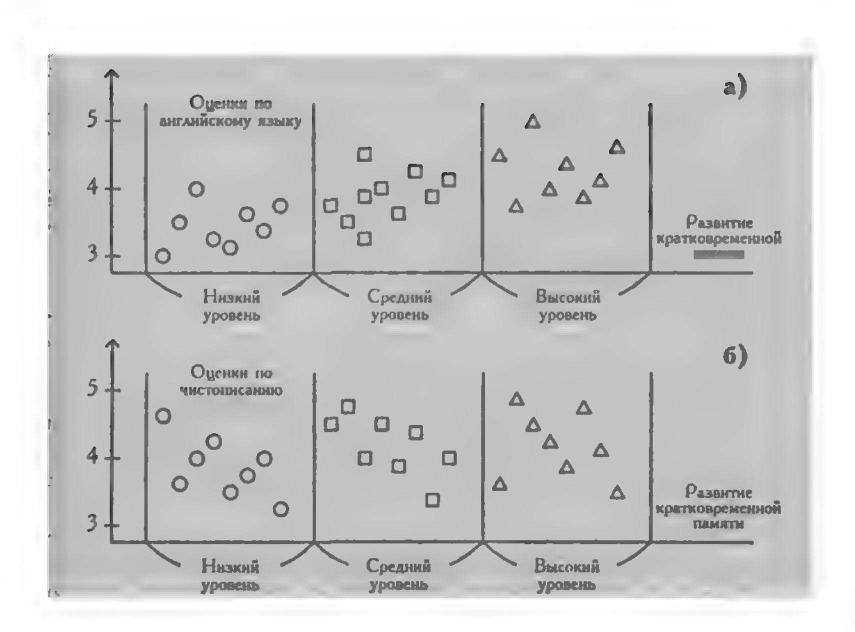


Рис. 7.1. Рассеивание индивидуальных средних оценок по английскому языку (а) и чистописанию (б) у учеников с низким, средним и высоким уровнями развития кратковременной памяти

Низкий, средний и высокий уровни развития кратковременной памяти можно рассматривать как градации фактора кратковременной памяти.

Нулевая гипотеза в дисперсионном анализе будет гласить, что средние величины исследуемого результативного признака во всех градациях одинаковы.

Альтернативная гипотеза будет утверждать, что средние величины результативного признака в разных градациях исследуемого фактора различны.

В зарубежных руководствах чаще говорят о переменных, действующих в разных условиях, а не о факторах и их градациях (Greene J., D'Olivera M, 1982, p. 91-93).

Дело в том, что градация подразумевает ступень, стадию, уровень развития. Говоря о градациях фактора, мы явно или неявно подразумеваем, что сила его возрастает при переходе от градации к градации. Между тем, схема дисперсионного анализа применима и в тех случаях, когда градации фактора представляют собой номинативную шкалу, то есть отличаются лишь качественно. Например, градациями фактора могут быть: параллельные формы экспериментальных заданий; цвет окраски стимулов; жанр музыкальных произведений, сопровождающих процесс работы; традиционные или специально подобранные православные тексты в сеансах аутогенной тренировки; разные формы заболевания; разные экспериментаторы; разные психотерапевты и т. д.

Если градации фактора различаются лишь качественно, их лучше называть *условиями* действия фактора или переменной. Например, действие аутогенной тренировки при условии использования текстов православных молитв или эффективность психокоррёкционных воздействий при разных формах хронических заболеваний у детей[3].

Экспериментальные данные, представленные по градациям фактора, называются дисперсионным комплексом. Данные, относящиеся к отдельным градациям - ячейками комплекса.

Дисперсионный анализ позволяет нам констатировать изменение признака, но при этом не указывает *направление* этих изменений. Нам необходимо специально графически представлять полученные данные по градациям фактора, чтобы получить наглядное представление о направлении изменений.

Подобного рода задачи, как мы помним, позволяют решать непараметрические методы сравнения выборок или условий измерения, а именно критерий Н. Крускала-Уоллиса и критерий χ^2_r Фридмана (см. параграфы 2.4 и 3.4). Однако это касается только тех задач, в которых исследуется действие *одного* фактора, или одной переменной. Задачи однофакторного дисперсионного анализа, действительно, могут эффективным образом

решаться с помощью непараметрических методов. Метод дисперсионного анализа становится незаменимым только когда мы исследуем одновременное действие *двух* (или более) факторов, поскольку он позволяет выявить *взаимодействие* факторов в их влиянии на один и тот же результативный признак. Именно эти возможности двухфакторного дисперсионного анализа послужили причиной, по которой изложение этого метода включено в настоящее руководство.

Несмотря на то, что нас интересует прежде всего двухфакторный дисперсионный анализ, который нельзя заменить другими методами, начнем рассмотрение мы с однофакторного дисперсионного анализа: во-первых, для того, чтобы выдержать определенную последовательность и логику в изложении; во-вторых, для того, чтобы на реальном примере продемонстрировать возможность замены этого метода непараметрическими методами.

См. исследование Е. Б. Кулевой, 1991. См. исследование Н.В.Корольковой, 1994.

Итак, начнем рассмотрение дисперсионного анализа с простейшего случая, когда исследуется действие только *одной* переменной (одного фактора). Исследователя интересует, как изменяется определенный признак в разных условиях действия этой переменной. Например, как изменяется время решения задачи при разных условиях мотивации испытуемых (низкой, средней, высокой) или при разных способах предъявления задачи (устно, письменно, в виде текста с графиками и иллюстрациями), в разных условиях работы с задачей (в одиночестве, в одной комнате с экспериментатором, в одной комнате с экспериментатором и другими испытуемыми) и т.п. В первом случае переменной, влияние которой исследуется, является мотивация, во втором - степень наглядности, в третьем - фактор публичности.

Преимущество однофакторного дисперсионного анализа по сравнению с непараметрическими методами Н Крускала-Уоллиса и χ^2_r Фридмана - неограниченность в объемах выборок. Ограничения дисперсионного анализа достаточно условны. Например, требование нормальности распределения признака можно обойти по крайней мере двумя путями: при слишком скошенном, островершинном или плосковершинном распределении можно, во-первых, нормализовать данные, а во-вторых... просто вообще по этому поводу "не волноваться", как советуют, например, A.K. Kurtz и S.T. Mayo (1979, p.417).

7.2. Подготовка данных к дисперсионному анализу

1) Создание комплексов

Лучше всего для каждого испытуемого создать отдельную карточку, куда были бы занесены данные по всем исследованным признакам. Дело в том, что в процессе анализа у исследователя могут измениться гипотезы. Потребуется создавать, быть может, не один, а множество дисперсионных комплексов, различающихся как по факторам, так и по результативным признакам. Карточки помогут нам быстро создавать новые дисперсионные комплексы. Благодаря карточкам мы сразу увидим, равномерно ли распределяются данные по градациям в случае, если за фактор мы решили принять один из исследованных психологических признаков. С помощью карточек мы можем помочь себе выделить три, четыре или более градаций этого фактора, например, уровни мотивации, настойчивости, креативности и др.

2) Уравновешивание комплексов

Комплекс, в котором каждая ячейка представлена одинаковым количеством наблюдений, называется равномерным. Равномерность комплекса позволяет нам обойти требование равенства дисперсий в каждой из ячеек комплекса (Шеффе Г., 1980).

Равномерные комплексы позволяют также избежать значительных трудностей, которые неизбежно возникают при обсчете неравномерных, или неортогональных, комплексов. В настоящем руководстве приведены алгоритмы расчета лишь для равномерных комплексов. С методами обсчета неравномерных комплексов можно ознакомиться у НА. Плохинского (1970), Г.В. Суходольского (1972), Г. Шеффе (1980).

В случае, если в разных градациях комплекса оказалось неравное количество наблюдений, необходимо отсеять некоторые из них. Если в комплексе со связанными выборками кто-либо из испытуемых не был подвергнут одному из условий действия переменной (градаций фактора), то его данные исключаются. Если же комплекс включает независимые выборки, каждая из которых была подвергнута определенному условию воздействия (градации фактора), то "лишние" испытуемые в какой-либо из ячеек комплекса отсеиваются путем случайного выбора необходимого количества карточек.

3) Проверка нормальности распределения результативного признака.

Дисперсионный анализ относится к группе параметрических методов и поэтому его следует применять только тогда, когда известно или доказано, что распределение признака является нормальным (Суходольский Г.В., 1972; Шеффе Г., 1980 и др.). Строго говоря, перед тем, как применять дисперсионный анализ, мы должны убедиться в нормальности распределения результативного признака. Нормальность распределения результативного признака можно проверить путем расчета показателей асимметрии и эксцесса и сопоставления их с критическими значениями (Пустыльник Е.И., 1968* Плохинский Н.А., 1970 и др.).

Произведем необходимые расчеты на примере параграфа 8.3, в котором анализируется длительность мышечного волевого усилия.

Действовать будем по следующему алгоритму:

а) определим показатели асимметрии и эксцесса по формулам Н.А. Плохинского и сопоставим их с критическими значениями, указанными Н.А. Плохинским;

б) рассчитаем критические значения показателей асимметрии и эксцесса по формулам Е.И. Пустыльника и сопоставим с ними эмпирические значения;

в) если эмпирические значения показателей окажутся ниже критических, сделаем вывод о том, что распределение признака не отличается от нормального.

Таблица ***7.1***

Вычисление показателей асимметрии и эксцесса по показателю длительности попыток решения анаграмм

№	x_i	$(x_i \ \bar{x})$	$(x_i \ \bar{x})^2$	$(x_i \ \bar{x})^3$	$(x_i \ \bar{x})^4$
1	11	0,94	0,884	0.831	0,781
2	13	2,94	8,644	25,412	74,712
3	12	1.94	3,764	7,301	14,165
4	9	-1,06	1,124	-1,191	1,262
5	10	-0.06	0,004	-0,000	0,000
6	11	0,94	0,884	0,831	0,781
7	8	-2,06	4,244	-8.742	18,009
8	10	-0,06	0,004	-0,000	0,000
9	15	4,94	24,404	120,554	595,536
10	14	3,94	15,524	61,163	240,982
И	8	-2,06	4,244	-8,742	18,009
12	7	-3.06	9,364	-28,653	87,677
13	10	-0.06	0,004	-0,000	0,000
14	10	-0,06	0.004	-0,000	0,000
15	5	-5,06	25,604	-129,554	655,544
16	8	-2,06	4,244	-8,742	18,009
Суммы	161		102,944	30,468	1725,467

Для расчетов в Табл. 7.1 необходимо сначала определить среднюю арифметическую по формуле:

$$\bar{x} = \frac{\sum x_i}{n}$$

где x_i - каждое наблюдаемое значение признака;

n - количество наблюдений. В данном случае:

$$\bar{x} = \frac{161}{16} = 10,06$$

Стандартное отклонение (сигма) вычисляется по формуле:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

где x_i - каждое наблюдаемое значение признака; $\bar{x}$ - среднее значение (среднее арифметическое); n - количество наблюдений. В данном случае:

$$\sigma = \sqrt{\frac{102{,}944}{16-1}} = \sqrt{6{,}893} = 2{,}62$$

Показатели асимметрии и эксцесса с их ошибками репрезентативности определяются по следующим формулам:

$$A = \frac{\sum(x_i - \bar{x})^3}{n \cdot \sigma^3}$$

$$m_A = \sqrt{\frac{6}{n}}$$

$$E = \frac{\sum(x_i - \bar{x})^4}{n \cdot \sigma^4} - 3$$

$$m_E = 2 \cdot \sqrt{\frac{6}{n}}$$

где $(x_i - \bar{x})$ - центральные отклонения;

σ - стандартное отклонение;

n - количество испытуемых. В данном случае:

$$A = \frac{+30{,}468}{16 \cdot 2{,}62^3} = +0{,}106$$

$$m_A = \sqrt{\frac{6}{16}} = 0{,}61$$

$$E = \frac{1725{,}467}{16 \cdot 2{,}62^4} - 3 = -0{,}711$$

$$m_E = 2 \cdot \sqrt{\frac{6}{16}} = 1{,}22$$

Показатели асимметрии и эксцесса свидетельствуют о достоверном отличии эмпирических распределений от нормального в том случае, если они превышают по абсолютной величине свою ошибку репрезентативности в 3 и более раз:

$$t_A = \frac{|A|}{m_A} \geq 3$$

$$t_E = \frac{|E|}{m_E} \geq 3$$

В данном случае:

$$t_A = \frac{|0{,}106|}{0.61} = 0{,}174$$

$$t_E = \frac{|-0{,}711|}{1{,}22} = 0{,}583$$

Мы видим, что оба показателя не превышают в три раза свою ошибку репрезентативности, из чего мы можем заключить, что распределение данного признака не отличается от нормального.

Теперь произведем проверку по формулам Е.И. Пустыльника. Рассчитаем критические значения для показателей A и E:

$$A_{кр} = 3 \cdot \sqrt{\frac{6 \cdot (n-1)}{(n+1) \cdot (n+3)}}$$

$$E_{кр} = 5 \cdot \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n+1)^2 \cdot (n+3) \cdot (n+5)}}$$

где n - количество наблюдений.

В данном случае:

$$A_{кр} = 3 \cdot \sqrt{\frac{6 \cdot (16-1)}{(16+1) \cdot (16+3)}} = 3 \cdot \sqrt{\frac{90}{323}} = 1{,}58$$

$$E_{кр} = 5 \cdot \sqrt{\frac{24 \cdot 16 \cdot (16-2) \cdot (16-3)}{(16+1)^2 \cdot (16+3) \cdot (16+5)}} = 5 \cdot \sqrt{\frac{69888}{115311}} = 3{,}89$$

$A_{эмп}=0{,}106$

$A_{эмп}<A_{кр}$

$E_{эмп}=-0{,}711$

$E_{эмп}<E_{кр}$

Итак, оба варианта проверки, по Н.А. Плохинскому и по Е.И. Пустыльнику, дают один и тот же результат: распределение результативного признака в данном примере не отличается от нормального распределения.

Можно выбрать любой из двух предложенных вариантов проверки и придерживаться его. При больших объемах выборки, по-видимому, стоит производить расчет первичных статистик (оценок параметров) на ЭВМ.

4) Преобразование эмпирических данных с целью упрощения расчетов

Н.А. Плохинский указывает на возможность следующих преобразований:

1) все наблюдаемые значения можно разделить на одно и то же число *k,* например перевести показатели из миллиметров в сантиметры и т.п.;

2) все наблюдаемые значения можно умножить на одно и то же число *k,* например для того, чтобы избавиться от дробных значений;

3) от всех наблюдаемых значений можно отнять одно и то же число *A,* например наименьшее значение;

4) можно сделать двойное преобразование: из каждого значения вычесть число *A,* а полученный результат разделить на другое число *k.*

При всех этих преобразованиях результативного признака показатели соотношения дисперсий получаются точными и не требуют никаких поправок.

Средние величины изменяются, но их можно восстановить, умножая среднюю величину на число *k* или деля ее на *k* (варианты 1 и *2)* или прибавляя к средней число *A* (вариант 3) и т. п. Стандартное отклонение изменяется только при введении множителя или делителя; полученный результат затем придется либо разделить на число *к,* либо умножить на него (Плохинский Н.А.,1964, с.34-36; Плохинский Н.А., 1970, с.71-72).

В последующих трех параграфах будет рассмотрен метод одно-факторного анализа в двух вариантах:

а) для дисперсионных комплексов, представляющих данные одной и той же выборки испытуемых, подвергнутой влиянию разных условий (разных градаций фактора);

б) для дисперсионных комплексов, в которых влиянию разных условий (градаций фактора) были подвергнуты разные выборки испытуемых.

Первый вариант называется однофакторным дисперсионным анализом для связанных выборок, второй - для несвязанных выборок.

Все предложенные алгоритмы расчетов предназначены для равномерных

комплексов, где в каждой ячейке представлено одинаковое | число наблюдений.

7.3. Однофакторный дисперсионный анализ для несвязанных выборок

Назначение метода

Метод однофакторного дисперсионного анализа применяется в тех |случаях, когда исследуются изменения результативного признака под [влиянием изменяющихся условий или градаций какого-либо фактора. В данном варианте метода влиянию каждой из градаций фактора подвергаются *разные* выборки испытуемых. Градаций фактора должно быть не менее трех[4].

Непараметрическим вариантом этого вида анализа является критерий Н Крускала-Уоллиса.

Описание метода

Работу начинаем с того, что представляем полученные данные в виде столбцов индивидуальных значений. Каждый из столбцов соответствует тому или иному из изучаемых условий (см. Табл. 7.2).

После этого нам нужно просуммировать индивидуальные значения по столбцам и суммы возвести в квадрат.

Суть метода состоит в том, чтобы сопоставить сумму этих возведенных в квадрат сумм с суммой квадратов всех значений, полученных во всем эксперименте.

[4] Градаций может быть и две, но в этом случае мы не сможем установить нелинейных зависимостей и более разумным представляется использование более простых критериев (см. главы 2 и 3).

Гипотезы

H_0: Различия между градациями фактора (разными условиями) являются не более выраженными, чем случайные различия внутри каждой группы.

H_1: Различия между градациями фактора (разными условиями) являются более выраженными, чем случайные различия внутри каждой группы.

Графическое представление метода для несвязанных выборок

На Рис. 7.2 показана кривая изменения объема воспроизведения слов при разной скорости их предъявления (см. Пример). Метод дисперсионного анализа позволяет определить, что перевешивает - тенденция, выраженная этой кривой, или вариативность признака внутри групп, которая на графике схематически изображена в виде диапазонов изменения признака от минимального значения к максимальному значению в каждой группе.

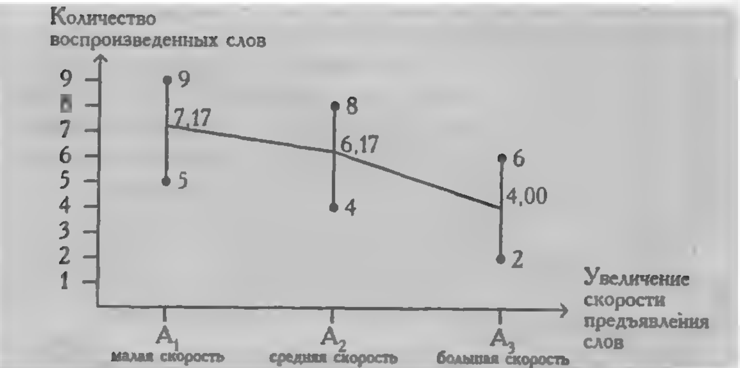


Рис. 7.2. Кривая изменения объема воспроизведения при повышении скорости предъявления слов; по каждому условию показаны диапазоны изменения признака (по данным Greene J., D'Olivera M., 1989)

Ограничения метода однофакторного дисперсионного анализа для несвязанных выборок

1. Однофакторный дисперсионный анализ требует не менее трех градаций фактора и не менее двух испытуемых в каждой градации.

2. Должно соблюдаться правило равенства дисперсий в каждой ячейке дисперсионного комплекса. Условие равенства дисперсий выполняется при

использовании предлагаемой схемы расчета за счет выравнивания

количества наблюдений в каждом из условий (градаций). Правомерность этого методического приема была обоснована Г.Шеффе (1980).

3. Результативный признак должен быть нормально распределен в исследуемой выборке.

Правда, обычно не указывается, идет ли речь о распределении I признака во всей обследованной выборке или в той ее части, которая [составляет дисперсионный комплекс.

Характерно, что зарубежные руководства, в общем ссылаясь на необходимость нормального распределения данных для дисперсионного анализа, при рассмотрении конкретных схем и примеров к этому вопросу уже не возвращаются и никаких данных о распределении признака в выборке в целом или в той ее части, которая составляет дисперсионный комплекс, не приводят (см; McCall R., 1970; Welkowitz J., Ewen R.B., | Cohen J., 1982; Greene J., D'Olivera M., 1989).

Рассмотрим схему дисперсионного однофакторного анализа для несвязанных выборок, предлагаемую в руководстве J.Greene, M.D'Olivera (1989) с использованием примера этих авторов.

Пример

Три различные группы из шести испытуемых получили списки из десяти слов. Первой группе слова предъявлялись с низкой скоростью -1 слово в 5 секунд, второй группе со средней скоростью - 1 слово в 2 секунды, и третьей группе с большой скоростью - 1 слово в секунду. Было предсказано, что показатели воспроизведения будут зависеть от скорости предъявления слов. Результаты представлены в Табл. 7.2.

Таблица 7.2 Количество воспроизведенных слов (по: J.Greene, M.D'Olivera, 1989,p.99)

№ испытуемого	Группа 1: низкая скорость	Группа 2: средняя скорость	Группа 3: высокая скорость
1	8	7	4
2	7	8	5
3	9	5	3
4	5	4	6
5	6	6	2
6	8	7	4
Суммы	43	37	24
Средние	7,17	6,17	4,00
Общая сумма		104	

Поскольку сопоставляются разные группы, любые различия в показателях между разными условиями предъявления слов - это в то же время различия между группами испытуемых. Однако всякие различия между испытуемыми *внутри* каждой группы объясняются какими-то Другими, не относящимися к делу переменными, будь то индивидуальные различия между отдельными испытуемыми или неконтролируемые факторы, заставляющие их реагировать различным образом. Критерий F позволяет проверить гипотезы:

H_0: Различия в объеме воспроизведения слов *между* группами являются не более выраженными, чем случайные различия *внутри* каждой группы.

H_1: Различия в объеме воспроизведения слов *между* группами являются более выраженными, чем случайные различия *внутри* каждой группы. Используя экспериментальные значения, представленные в Табл. 7.2, установим некоторые величины, которые будут необходимы для расчета критерия F.

Таблица 7.3 Расчет основных величин для однофакторного дисперсионного анализа

Обозначение	Расшифровка обозначения	Экспериментальные значения
T_c	суммы индивидуальных значений по каждому из условий	43; 37; 24
$\Sigma(T^2_c)$	сумма квадратов суммарных значений по каждому из условий	$\Sigma(T^2_c)=43^2+37^2+24^2$
c	количество условий (градаций фактора)	$c=3$
n	количество испытуемых в каждой группе (в каждом из условий)	$n=6$
N	общее количество индивидуальных значений	$N=18$
$(\Sigma x_i)^2$	квадрат общей суммы индивидуальных значений	$(\Sigma x_i)^2=104^2$
$\frac{(\Sigma x_i)^2}{N}$	константа, которую нужно вычесть из каждой суммы квадратов	$\frac{(\Sigma x_i)^2}{N}=\frac{104^2}{18}$
x_i	каждое индивидуальное значение	
$\Sigma(x_i^2)$	сумма квадратов индивидуальных значений	

Отметим разницу между $\sum(x_i^2)$, в которой все индивидуальные значения сначала возводятся в квадрат, а потом суммируются, и $(\sum x_i)^2$ где индивидуальные значения сначала суммируются для получения об- j щей суммы, а потом уже эта сумма возводится в квадрат.

Последовательность расчетов представлена в Табл. 7.4.

Часто встречающееся в этой и последующих таблицах обозначение SS - сокращение от "суммы квадратов" (sum of squares). Это сокращение чаще всего используется в переводных источниках (см., например: Гласе Дж., Стенли Дж., 1976).

$SS_{факт}$ означает вариативность признака, обусловленную действием исследуемого фактора; $SS_{общ}$ - общую вариативность признака; $SS_{СА}$ -вариативность, обусловленную неучтенными факторами, "случайную" или "остаточную" вариативность.

MS - "средний квадрат", или математическое ожидание суммы квадратов, усредненная величина соответствующих SS.

df - число степеней свободы, которое при рассмотрении непараметрических критериев мы обозначили греческой буквой *v*.

Таблица 7.4

Последовательность операций в однофакторном дисперсионном анализе для несвязанных выборок

Операция	Формула расчета	Расчет по экспериментальным данным
1.Подсчитать $SS_{факт}$	$SS_{факт}={}^1/_n\cdot\Sigma T^2_c-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_{факт}=(43^2+37^2+24^2)/6-104^2/18=31,44$
2.Подсчитать $SS_{общ}$	$SS_{общ}=\Sigma x^2_i-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_{общ}=8^2+7^2+9^2+5^2+6^2+8^2+7^2+8^2+5^2+4^2+6^2+7^2+4^2+5^2+3^2+6^2+2^2+4^2-104^2/18==63,11$
3.Подсчитать случайную (остаточную) величину $SS_{сл}$	$SS_{сл}=SS_{общ}-SS_{факт}$	$SS_{сл}=63,11-31,44=31,67$
4.Определить число степеней свободы	$df_{факт}=c-1$ $df_{общ}=N-1$ $df_{сл}=df_{общ}-df_{факт}$	$df_{факт}=3-1=2$ $df_{общ}=18-1=17$ $df_{сл}=17-2=15$
5.Разделить каждую SS на соответствующее число степеней свободы	$MS_{факт}=SS_{факт}/df_{факт}$ $MS_{сл}=SS_{сл}/df_{сл}$	$MS_{факт}=31,44/2=15,72$ $MS_{сл}=31,67/15=2,11$
6.Подсчитать значение $F_{эмп}$	$F_{эмп}=MS_{факт}/MS_{сл}$	$F_{эмп(2,15)}=15,72/2,11=7,45$
7.Определить критические значения F по Табл. XVII Приложения 1	для $df_1=2$ $df_2=15$	$F_{кр\ (2,15)}=\begin{cases}3,68\ (\rho\le0,05)\\6,36\ (\rho\le0,01)\end{cases}$
8.Сопоставить эмпирическое и критические значения F	При $F_{эмп}\ge F_{кр}$ H_0 отклоняется	$F_{эмп}>F_{кр}\rightarrow H_0$ отклоняется

Вывод: H_0 отклоняется. Принимается H_1. Различия в объеме [воспроизведения слов между группами являются более выраженными, [чем случайные различия внутри каждой группы (p<0,01). Итак, скорость предъявления слов влияет на объем их воспроизведения[5]. Вернемся к графику на Рис. 7.2. Мы видим, что, скорее всего, значимость различий объясняется тем, что показатель воспроизведения при самой высокой скорости

предъявления слов (условие 3) гораздо ниже соответствующих показателей при средней и низкой скорости.

7.4. Дисперсионный анализ для связанных выборок

Назначение метода

Метод дисперсионного анализа для связанных выборок применяется в тех случаях, когда исследуется влияние разных градаций фактора или разных условий на *одну и ту же* выборку испытуемых.

Градаций фактора должно быть не менее трех.

Непараметрический вариант этого вида анализа - критерий Фридмана χ^2_r.

Описание метода

В данном случае различия между испытуемыми - возможный самостоятельный источник различий. В схеме однофакторного анализа для несвязанных выборок различия между условиями в то же время отражали различия между испытуемыми. Теперь различия между условиями могут проявиться только вопреки различиям между испытуемыми.

Фактор индивидуальных различий может оказаться более значимым, чем фактор изменения экспериментальных условий. Поэтому нам необходимо учитывать еще одну величину - сумму квадратов сумм индивидуальных значений испытуемых.

---

[5] Г.В. Суходольским (1972) предложена формула расчета дисперсионного отношения, которая позволяет получить более строгий результат:

$F_{эмп}=(n \cdot MS_{факт}+MS_{сл})/MS_{сл}$

где *n* - среднее количество наблюдений в каждой градации.

В данном случае $F_{амп}=6{,}942$ ($p<0{,}01$). Эта величина действительно ниже, чем в цитируемом примере. Однако для первого знакомства с дисперсионным анализом исследователям, обрабатывающим свои данные самостоятельно, в практических целях достаточно использовать приведенный алгоритм расчетов, используемый и в большинстве других руководств (Плохинский Н.А., 1960; Венецкий И.Г., Кильдишев ГС, 1968; Ивантер Э.В., Коросов А.В.; 1992, Kurtz A.K., Mayo S.T, 1979 и др.).

Графическое представление метода

На Рис. 7.3 представлена кривая Изменения времени решения анаграмм разной длины: четырехбуквенной, пятибуквенной и шестибуквенной. Однофакторный дисперсионный анализ для связанных выборок позволит определить, что перевешивает - тенденция, выраженная этой кривой, или индивидуальные различия, диапазон которых представлен на графике в виде вертикальных линий – от минимального до максимального значения.

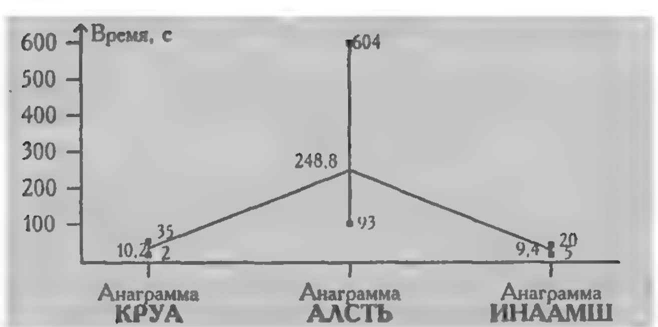


Рис. 7.3. Изменение времени работы над разными анаграммами у пяти испытуемых; вертикальными линиями отображены диапазоны изменчивости признака в разных условиях от минимального значения (снизу) до максимального значения (сверху)

Ограничения метода дисперсионного анализа для связанных выборок

1. Дисперсионный анализ для связанных выборок требует не менее *трех* градаций фактора и не менее *двух* испытуемых, подвергшихся воздействию каждой из градаций фактора.
2. Должно соблюдаться правило равенства дисперсий в каждой ячейке комплекса.

Это условие косвенно выполняется за счет одинакового количества наблюдений в каждой ячейке комплекса. Предлагаемая схема расчета ориентирована только на такие равномерные комплексы.

3. Результативный признак должен быть нормально распределен в исследуемой выборке.

В приводимом ниже примере показатели асимметрии и эксцесса составляют:

A=2,18;
m_A=0,632;
t_A=2,18/0,632=3,45;
E=4,17;
m_E=1,264;
t_E=4,17/1,264=3,30.

Таким образом, распределение показателей 5-ти, человек, составляющих дисперсионный комплекс, несколько отличается от нормального: $t_A>3$; $t_E>3$. Однако в целом по выборке распределение нормальное:

n=22; A=1,26; m_A=0,522 t_A=2,41<3; E=2,29; m_E=1,044; t_E=2,19<3.

По-видимому, необходимо удовлетвориться тем, что в выборке в целом результативный признак распределен нормально. Случайно отобранные 5 человек распределением своих оценок демонстрируют некоторое отклонение. Однако, если бы мы выбирали испытуемых таким образом, чтобы распределение их оценок подчинялось нормальному закону, это нарушило бы правило рандомизации - случайности отбора объектов без учета значений результативного признака при отборе (Плохинский Н.А., 1970).

Данные этого примера нам уже знакомы. Они использовались для иллюстрации непараметрического критерия Фридмана χ^2_r. Использование здесь этого же примера позволит нам сопоставить результаты, получаемые с помощью непараметрических и параметрических методов.

Пример

Группа из 5 испытуемых была обследована с помощью трех экспериментальных заданий, направленных на изучение интеллектуальной, настойчивости (Сидоренко Е. В., 1984). Каждому испытуемому индивидуально предъявлялись последовательно три одинаковые анаграммы: четырехбуквенная, пятибуквенная и шестибуквенная. Можно ли считать, что фактор длины анаграммы влияет на длительность попыток ее решения?

Сформулируем гипотезы. Наборов гипотез в данном случае два. *Набор А.*

$H_{0(A)}$: Различия в длительности попыток решения анаграмм разной длины являются не более выраженными, чем различия, обусловленные случайными причинами.

$H_{1(A)}$: Различия в длительности попыток решения анаграмм разной длины являются более выраженными, чем различия, обусловленные случайными причинами. *Набор Б.*

$H_{0(Б)}$: Индивидуальные различия между испытуемыми являются не более выраженными, чем различия, обусловленные случайными причинами.

$H_{1(Б)}$: Индивидуальные различия между испытуемыми являются более выраженными, чем различия, обусловленные случайными причинами.

Длительность попыток решения анаграмм (сек)

Таблица 7.5

Код имени испытуемого	Условие 1: четырехбуквенная анаграмма	Условие 2: пятибуквенная анаграмма	Условие 3; шестибуквенная анаграмма	Суммы по испытуемым
1. Л-в	5	235	7	247
2. П-о	7	604	20	631
3. К-в	2	93	5	100
4. Ю-ч	2	171	8	181
5. Р-о	35	141	7	183
Суммы по столбцам	51	1244	47	1342

Установим все промежуточные величины; необходимые для расчета критерия F.

Таблица 7.6

Расчет промежуточных величин для критерия F в примере об анаграммах

Обозначение	Расшифровка обозначения	Экспериментальное значение
T_c	суммы индивидуальных значений по каждому из условий (столбцов)	51; 1244; 47
$\sum T^2_c$	сумма квадратов суммарных значений по каждому из условий	$\sum T^2_c = 51^2 + 1244^2 + 47^2$
n	количество испытуемых	n=5
c	количество значений у каждого испытуемого (т. е. количество условий)	c=3
N	общее количество значений	N=15
T_n	суммы индивидуальных значений по каждому испытуемому	247; 631; 100; 181; 183
$\sum T^2_n$	сумма квадратов сумм индивидуальных значений по испытуемым	$247^2+631^2+100^2+181^2+183^2$
$(\sum x_i)^2$	квадрат общей суммы индивидуальных значений	$(\sum x_i)^2 = 1342^2$
$^1/_N * (\sum x_i)^2$	константа, которую нужно вычесть из каждой суммы квадратов	$^1/_N * (\sum x_i)^2 = {^1/_N} * 1342^2$
x_i	каждое индивидуальное значение	
$\sum x^2_i$	сумма квадратов индивидуальных значений	

Мы по-прежнему помним разницу между квадратом суммы и суммой квадратов Последовательность расчетов приведена в Табл. 7.7.

ГЛАВА 8 ДИСПЕРСИОННЫЙ ДВУХФАКТОРНЫЙ АНАЛИЗ

8.1. Обоснование задачи по оценке взаимодействия двух факторов

Двухфакторный дисперсионный анализ позволяет нам оценить не только влияние каждого из факторов в отдельности, но и их взаимодействие. Может оказаться, что одна переменная значимо действует на исследуемый признак только при малых (или, напротив, больших) значениях другой переменной. Например, повышение вознаграждения может повышать скорость решения задач у высокоинтеллектуальных испытуемых и понижать ее у низкоинтеллектуальных. Усиление наказания может снижать количество агрессивных реакций у девочек и повышать его у мальчиков. Или, скажем, внушение может влиять на младших школьников, но не влиять на подростков. Один фактор может "заморозить" или, напротив, "катализировать" действие другого.

В исследовании К.А. Harris и К.В. Morrow изучалась такая личностная черта, как доминантность взрослых мужчин и женщин: Авторы предполагали, что доминантность должна быть выше у людей, которые были первенцами в своих семьях, и ниже у средних и тем более младших детей. Оказалось, что влияние каждого из двух исследуемых факторов - пола и порядка рождения - незначимо, а взаимодействие факторов значимо (см. Рис. 8.1). У мужчин доминантность, как и предполагалось, с увеличением порядка рождения снижается, а у женщин, напротив, повышается. Авторы объясняют это двояко: тем, что младшие девочки в семьях могут пользоваться особым предпочтением остальных членов семьи или тем, что повышенной доминантностью они отвечают на свое подчиненное положение в детстве (Harris K.A., Morrow K.B.,g 1992).

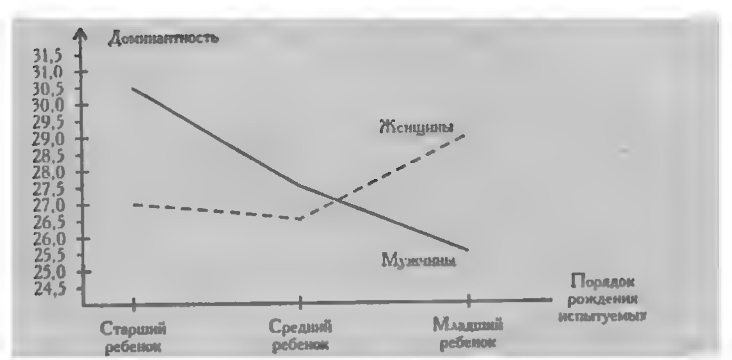


Рис. 3.1. Изменения показателей Доминантности (шкала Калифорнийского личностного опросника) в зависимости от порядка рождения у мужчин (сплошная линия) и женщин (пунктирная линия) (по: Harris А. К., Morrow К. В., 1992, р. 115)

Если нами установлено значимое взаимодействие факторов, то это зачастую важнее, чем действие каждого из факторов в отдельности. Некоторые исследователи предлагают вообще игнорировать в таких случаях "основные эффекты" каждого из взаимодействующих факторов и рассматривать только взаимодействие (McCall R., 1970, p. 250).

Специалист по возрастной и дифференциальной психологии знает, что "основных эффектов", или общих закономерностей, в действительности достаточно мало. Почти всегда требуется поправка на возраст испытуемых, их пол, профессиональную принадлежность, способ восприятия, тип энергетической мобилизации и т.п. К счастью, петербургская-ленинградская школа психологии благодаря, в первую очередь, Б.Г., Ананьеву, никогда не была "бесполой" или "вневозрастной" (см., например, Ананьев Б.Г., 1968). Именно поэтому дисперсионный анализ в большей степени отвечает ленинградскому дифференциально-психологическому подходу в экспериментальных исследованиях. Он помогает нам выявлять все более и более частные и точные закономерности и приближает нас к установлению закономерностей *индивидуальных стилей.*

Двухфакторный дисперсионный анализ предъявляет особые требования к

формированию комплексов. Комплекс должен представлять собой симметричную систему: каждой градации фактора А должно соответствовать одинаковое количество градаций фактора В. Например, для исследования A.K. Harris, K.B. Morrow (см. Рис. 8.1) это означает, что и среди мужчин должны были быть старшие, средние и младшие дети, и среди женщин должны быть старшие, средние и младшие дети, причем для равномерного комплекса необходимо, чтобы в каждой ячейке комплекса было одинаковое количество испытуемых. Понятно, конечно же, что это значительно усложняет исследование и требует тщательного предварительного планирования его.

Подробности работы лучше рассматривать на примерах, поэтому перейдем к моделям двухфакторного дисперсионного анализа: а) для несвязанных выборок; б) для связанных выборок.

8.2. Двухфакторный дисперсионный анализ для несвязанных выборок

Назначение метода

Данный вариант двухфакторного дисперсионного анализа применяется в тех случаях, когда исследуется одновременное действие *двух* факторов на *разные* выборки испытуемых, т. е. когда разные выборки, испытуемых оказываются под воздействием *разных сочетаний* двух факторов. Количество выборок определяется количеством ячеек дисперсионного комплекса.

Описание метода

Суть метода остается прежней, но в двухфакторном дисперсионном анализе мы можем проверить большее количество гипотез. Расчеты гораздо сложнее, чем в однофакторных комплексах.

Используемый в данном руководстве алгоритм расчетов предназначен только для равномерных комплексов. Если комплекс получился неравномерным, необходимо случайным образом отсеять несколько испытуемых.

Работу начинаем с построения специальной таблицы, отражающей весь дисперсионный комплекс. Подробности лучше сразу рассматривать на примере.

Пример

Рассмотрим пример из руководства J.Greene, M.D.'Olivera (1989).

Четырем группам испытуемых предъявлялись списки из 10 слов:

группе 1 - короткие слова с большой скоростью;

группе 2 - короткие слова с медленной скоростью;

группе 3 - длинные слова с большой скоростью;

группе 4 - длинные слова с медленной скоростью.

В каждой группе было по 4 испытуемых, всего N=16. Предсказывалось, что между факторами длины слов и скоростью их предъявления будет наблюдаться значимое *взаимодействие:* при большой скорости предъявления лучше будут запоминаться короткие слова, а при медленной скорости - длинные слова. Результаты экспериментов представлены в Табл. 8.1.

Таблица 8.1

Количество воспроизведенных слов при разной длине слов и разной скорости их предъявления (по J.Greene, M.D'Olivera, 1989)

Переменная (фактор) В скорость предъявления слов	Переменная (фактор) А - длина слов				Суммы по переменной В (T_B)
	A_1 - короткие слова		A_2 - длинные слова		
B_1 (большая скорость)	9867	30	5334	15	45
B_2 (малая скорость)	4335	15	7567	25	40
Суммы по переменной А (T_A)		45		40	85

Заметим, что в отечественных руководствах чаще предлагается другая, более

привычная для нас, форма таблиц для двухфакторных дисперсионных комплексов (Табл. 8.2). При такой форме легче "увидеть" комплекс в целом.

Таблица 8.2

Двухфакторный дисперсионный комплекс по оценке влияния фактора А (длина слов) и фактора В (скорость предъявления слов) на количество воспроизведенных слов

Градации фактора А	A_1 - короткие слова		A_2 – длинные слова	
Градации фактора В	B_1	B_2	B_1	B_2
	9 8 6 7	4 3 3 5	5 3 3 4	7 5 6 7
Суммы по ячейкам	30	15	15	25
Суммы по градациям фактора А	T_{A1}=45		T_{A2}=40.	
Суммы по градациям фактора В	T_{B1}==30+15=45		T_{B2}==15+25=40	

Как видим, при такой форме таблицы легче подсчитать суммы по ячейкам (в столбик), но труднее разобраться с суммами по градациям каждого из факторов. В данном случае оказалось, что они совпали:

$T_{A1}= T_{B1}$; $T_{A2}= T_{B2}$.

В дальнейшем при использовании алгоритма расчетов будем опираться на Табл. 8.1.

Сформулируем гипотезы. Это будут гипотезы, касающиеся влияния фактора А отдельно от фактора В (как бы при "усредненных" его значениях), гипотезы о влиянии фактора В отдельно от фактора А и гипотезы о влиянии взаимодействия градаций факторов А и В.

1 комплект гипотез

H_0: Различия в объеме воспроизведения слов, обусловленные действием фактора А, являются не более выраженными, чем случайные различия между показателями.

H_1: Различия в объеме воспроизведения слов, обусловленные действием фактора А, являются более выраженными, чем случайные различия между показателями.

2 комплект гипотез

H_0: Различия в объеме воспроизведения слов, обусловленные действием фактора В, являются не более выраженными, чем случайные различия между показателями.

H_1: Различия в объеме воспроизведения слов, обусловленные действием фактора В, являются более выраженными, чем случайные различия между показателями.

3 комплект гипотез

H_0: Влияние фактора А на объем воспроизведения слов одинаково при разных градациях фактора В, и наоборот.

H_1: Влияние фактора А на объем воспроизведения слов различно при разных градациях фактора В, и наоборот.

Используя экспериментальные значения, представленные в Табл. 8.1, установим некоторые величины, которые будут необходимы для расчета критериев F.

Таблица 8.3

Величины, необходимые для расчета критериев F в двухфакторном дисперсионном анализе для несвязанных выборок

Обозначение	Расшифровка обозначения	Экспериментальные значения
T_A	суммы по градациям фактора A	45; 40
ΣT^2_A	сумма квадратов этих сумм	$\Sigma T^2_A=45^2+40^2$
T_B	суммы по градациям фактора B	45; 40
ΣT^2_B	сумма квадратов этих сумм	$\Sigma T^2_B=45^2+40^2$
T_{AB}	суммы по "ячейкам"	30; 15; 15; 25
ΣT^2_{AB}	сумма квадратов этих сумм	$\Sigma T^2_{AB}=30^2+15^2+15^2+25^2$
n	количество испытуемых в каждой ячейке	$n=4$
a	количество градаций фактора A	$a=2$
b	количество градаций фактора B	$b=2$
N	общее количество индивидуальных значений	$N=16$
x_i	каждое индивидуальное значение	
Σx_i	общая сумма всех индивидуальных значений	$\Sigma x_i=85$
$(\Sigma x_i)^2$	квадрат общей суммы	$(\Sigma x_i)^2=85^2$
$^1/_N \cdot (\Sigma x_i)^2$	константа, которая вычитается из всех SS	$^1/_N \cdot (\Sigma x_i)^2=85^2/16$
Σx_i^2	сумма квадратов индивидуальных значений	

Напомним, что при подсчете $\sum x_i^2$ все индивидуальные значения сначала возводятся в квадрат, а потом суммируются, а при подсчете $(\sum x_i)^2$ все индивидуальные значения сначала суммируются, а затем их общая сумма возводится в квадрат.

Последовательность расчетов представлена в Табл. 8.4.

Таблица 8.4

Последовательность операций в двухфакторном дисперсионном анализе для несвязанных выборок

	Операция	Формула расчета	Расчет по экспериментальным данным Примера
1	Подсчитать SS_A	$SS_A={}^1/_{n\cdot b}\cdot\Sigma T^2_A-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_A=(45^2+40^2)/(4\cdot2)-85^2/16=$ $=3625/8-7225/16=1,56$
2	Подсчитать SS_B	$SS_B={}^1/_{n\cdot a}\cdot\Sigma T^2_B-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_B=(45^2+40^2)/(4\cdot2)-85^2/16=$ $=3625/8-7225/16=1,56$
3	Подсчитать SS_{AB}	$SS_{AB}={}^1/_n\cdot\Sigma T^2_{AB}-{}^1/_N\cdot(\Sigma x_i)^2-$ $-SS_A-SS_B$	$SS_{AB}=(30^2+15^2+15^2+25^2)/4-$ $-85^2/16-1,56-1,56=1975/4-$ $-7225/16-3,12=39,06$
4	Подсчитать $SS_{общ}$	$SS_{общ}=\Sigma x_i^2-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_{общ}=9^2+8^2+6^2+7^2+4^2+3^2+3^2+$ $+5^2+5^2+3^2+3^2+4^2+7^2+5^2+6^2+7^2-$ $-85^2/16=55,43$
5	Подсчитать SS сл	$SS_{сл}=SS_{общ}-SS_A-SS_B-SS_{AB}$	$SS_{сл}=55,43-1,56-1,56-39,06=13,25$
6	Подсчитать число степеней свободы	$df_A=a-1$ $df_B=b-1$ $df_{AB}=df_A\cdot df_B$ $df_{общ}=N-1$ $df_{сл}=df_{общ}-df_A-df_B-df_{AB}$	$df_A=2-1=1$ $df_B=2-1=1$ $df_{AB}=1\cdot1=1$ $df_{общ}=16-1=15$ $df_{сл}=15-1-1-1=12$
7	Разделить каждую SS на соответствующее число степеней свободы	$MS_A=SS_A/df_A$ $MS_B=SS_B/df_B$ $MS_{AB}=SS_{AB}/df_{AB}$ $MS_{сл}=SS_{сл}/df_{сл}$	$MS_A=1,56/1=1,56$ $MS_B=1,56/1=1,56$ $MS_{AB}=39,06/1=39,06$ $MS_{сл}=13,25/12=1,104$
8	Подсчитать значения F, определив df_1 по числителю и df_2 по знаменателю	$F_A=MS_A/MS_{сл}$ $F_B=MS_B/MS_{сл}$ $F_{AB}=MS_{AB}/MS_{сл}$	$F_{A\,(1,12)}=1,56/1,104=1,41$ $F_{B\,(1,12)}=1,56/1,104=1,41$ $F_{AB\,(1,12)}=39,06/1,104=35,38$
9	Определить критические значения F по Табл. XVII Приложения 1	На точке пересечения $df_1=1$, $df_2=12$	$F_{кр\,(1,12)}=\begin{cases}4,75 & (\rho\le0,05)\\ 9,33 & (\rho\le0,01)\end{cases}$
10	Сопоставить эмпирические значения F с критическими	При $F_{эмп}<F_{кр}$ H_0 принимается При $F_{эмп}\ge F_{кр}$ H_0 отвергается	$F_{эмпA}<F_{кр}$ → H_0 принимается $F_{эмпB}<F_{кр}$ → H_0 принимается $F_{эмпAB}>F_{кр}$ → H_0 отклоняется

Вывод: Но принимается в комплектах гипотез 1 и 2. Различия в объеме воспроизведения слов, обусловленные в отдельности факторами A и B, не являются более выраженными, чем случайные различия между показателями. H_0 отвергается для взаимодействия факторов (3 комплект). Принимается H_1. Влияние фактора A на объем воспроизведения слов различно при разных градациях фактора B, и наоборот ($p\leq0,01$).

Итак, оказывается, что факторы длины слов и скорости их предъявления в отдельности не оказывают значимого действия на объем воспроизведения. Значимым оказывается именно взаимодействие факторов: короткие слова лучше запоминаются при быстрой скорости предъявления, а длинные - при медленной скорости предъявления (см. Рис. 8.2). Таким образом, предположение, высказанное авторами, нашло статистически значимое подтверждение ($p\leq0,001$).

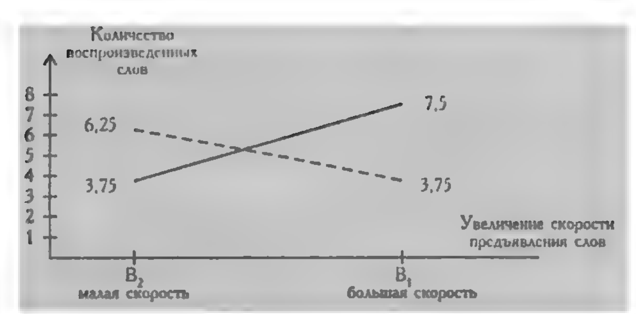


Рис. 8.2. Кривые изменения объема воспроизведения при повышении скорости предъявления коротких (сплошная линия) и длинных слов (пунктирная линия)

Ограничения двухфакторного дисперсионного анализа для несвязанных выборок

1. У каждого фактора должно быть не менее двух градаций.
2. В каждой ячейке комплекса должно быть не менее двух наблюдаемых значений для выявления взаимодействия градаций.
3. Количества значений во всех ячейках комплекса должны быть равны для обеспечения равенства дисперсий в ячейках комплекса и для использования приведенного выше алгоритма расчетов; для неравномерных комплексов можно использовать алгоритмы Н.А. Плохинского (1970).
4. Комплекс должен представлять собой симметричную систему: каждой градации фактора А должно соответствовать одинаковое количество градаций фактора В.
5. Результативный признак должен быть нормально распределен в исследуемой выборке, в противном случае значимые различия будет выявить гораздо труднее и применение метода будет не вполне корректным.
6. Факторы должны быть независимыми. В рассмотренном примере скорость предъявления слов и их длина - внешне независимые факторы. В других случаях независимость факторов может быть подтверждена отсутствием корреляционной связи между переменными, выступающими в качестве факторов.

8.3. Двухфакторный дисперсионный анализ для связанных выборок

Назначение метода

Данный вариант двухфакторного дисперсионного анализа применяется в тех случаях, когда исследуется действие *двух* факторов на *од-ну и ту же* выборку испытуемых.

Описание метода

Допустим, мы измерили одни и те же показатели у одних и тех же испытуемых несколько раз - в разное время, в разных условиях, с помощью параллельных форм методики и т. п., и нам необходимо провести множественное сравнение показателей, изменяющихся при переходе от условия к условию. Критерий L Пейджа для анализа тенденций изменения признака и критерий χ^2_r Фридмана неприменимы, так как необходимо определить тенденцию изменения признака под влиянием двух факторов одновременно. Это позволяет сделать только дисперсионный анализ.

Фактически в данной модели дисперсионного двухфакторного анализа проверяются 4 гипотезы: о влиянии фактора А, о влиянии фактора В, о влиянии взаимодействия факторов А и В и о влиянии фактора индивидуальных различий.

В данном варианте дисперсионного анализа нам потребуются две рабочие таблицы, которые позволят рассчитывать сумму по разным комбинациям ячеек комплекса. Рассмотрим это на примере, являющемся продолжением примера из п. 3.3.

Пример

В выборке курсантов военного училища (юноши в возрасте от 18 до 20 лет) измерялась способность к удержанию физического волевого усилия на динамометре. В первый день эксперимента у них, наряду с другими показателями, измерялась мышечная

сила каждой из рук. На второй день эксперимента им предлагалось выдерживать на динамометре мышечное усилие, равное $^1/_2$ максимальной мышечной силы данной руки. На третий день эксперимента испытуемым предлагалось проделать то же самое в парном соревновании на глазах у всей группы. Пары соревнующихся были подобраны таким образом, чтобы сила обеих рук у них примерно совпадала. Результаты экспериментов представлены в Табл. 8.5. Можно ли считать, что фактор соревнования в группе каким-то образом влияет на продолжительность удержания усилия? Подтверждается ли предположение о том, что правая рука более "социальна"?

Таблица 8.5

Длительность удержания усилия (сек/10) на динамометре правой и левой руками в разных условиях измерения (n=4)

Код имени испытуемого	Наедине с экспериментатором (A_1)		В группе сокурсников (A_2)	
	Правая рука	Левая рука	Правая рука	Левая рука
1 Л-в	11	10	15	10
2 С-с	13	11	14	10
3 С-в	12	8	8	5
4 К-в	9	10	7	8

Заметим, что единицы измерения в Табл. 8.5 - это секунды, но в каждом случае количество секунд уменьшено в 10 раз. Это законный способ преобразования индивидуальных значений, направленный на облегчение расчетов. Для того, чтобы не оперировать трехзначными числами, мы можем разделить их на какую-либо константную величину или уменьшить их на какую-либо константную величину (подробнее см п. 7.2).

Преобразуем таблицу индивидуальных значений в две рабочие таблицы двухфакторного дисперсионного комплекса для связанных выборок (Табл. 8.6 и 8.7). Мы видим, что здесь приведены суммы индивидуальных значений отдельно по градациям фактора А (вне группы - в группе) и по градациям фактора В (правая рука - левая рука), по сочетаниям градаций A_1B_1, A_1B_2, A_2B_1, A_2B_2, а также суммы всех индивидуальных значений каждого испытуемого и общие суммы.

Таблица 8.6 Двухфакторный дисперсионный комплекс по оценке влияния фактора А (вне группы - в группе) и фактора В (правая - левая рука) на длительность удержания физического волевого усилия (сек/10) - вариант I

Код имени испытуемого	A_1 - вне группы			A_2 - в группе			Индивидуальные суммы всех 4-х значений
	B_1	B_2	Индивидуальные суммы по A_1 (B_1+B_2)	B_1	B_2	Индивидуальные суммы по A_2 (B_1+B_2)	
1. Л-в	11	10	21	15	10	25	46
2. С-с	13	11	24	14	10	24	48
3. С-в	12	8	20	8	5	13	33
4. К-в	9	10	19	7	8	15	34
Суммы по ячейкам	45	39		44	33		
Суммы по градациям At и A?			84			77	
Общая сумма							161

Таблица 8.7

Двухфакторный дисперсионный комплекс по оценке влияния факторов А и В на длительность физического волевого усилия (сек/10) - вариант II

Код имени	B_1 – правая рука	B_2 – левая рука	Индивидуальные

испытуемого	A_1	A_2	Индивидуальные суммы по B_1 (A_1+A_2)	A_1	A_2	Индивидуальные суммы по B_2 (A_1+A_2)	суммы всех 4-х значений
1. Л-в	11	15	26	10	10	20	46
2. С-с	13	14	27	11	10	21	48
3. С-в	12	8	20	8	5	13	33
4. К-в	9	7	16	10	8	18	34
Суммы по ячейкам	45	44		39	33		
Суммы по градациям At и A?			89			72	
Общая сумма							161

Мы видим, что в Табл. 8.7 фактически только две ячейки комплекса поменялись местами: A_1B_2 и A_2B_1. Это позволяет нам с большей легкостью подсчитать суммы по градациям B_1 и B_2. Если бы 'мы пользовались только Табл. 8.6, то нам пришлось бы подсчитывать их "через столбец" и, кроме того, трудно было бы их куда-то подходящим образом записать. В дальнейшем при расчетах мы всякий раз будем указывать, к какой таблице лучше обратиться для извлечения нужных сумм, первой (I) или второй (II).

Установим некоторые величины, которые будут необходимы для расчёта критериев F.

Таблица 8.8

Величины, необходимые для расчета критериев F в двухфакторном дисперсионном анализе для связанных выборок

Обозначения	Расшифровка обозначения	Экспериментальные значения
$T_А$	суммы по градациям фактора А (I)	84; 77
$\Sigma T^2_А$	сумма квадратов этих сумм	$\Sigma T^2_А=84^2+77^2=12985$
$T_В$	суммы по градациям фактора В (II)	89; 72
$\Sigma T^2_В$	сумма квадратов этих сумм	$\Sigma T^2_В=89^2+72^2=13105$
$T_И$	индивидуальные суммы по 4-м значениям испытуемого (I или II)	46; 48; 33; 34
$\Sigma T^2_И$	сумма квадратов индивидуальных сумм по 4-м значениям	$\Sigma T^2_И=46^2+48^2+33^2+34^2=6665$
$\Sigma T^2_{АВ}$	сумма квадратов сумм по ячейкам (I, третья строка снизу)	$\Sigma T^2_{АВ}=45^2+39^2+44^2+33^2=6571$
$\Sigma T^2_{АИ}$	сумма квадратов индивидуальных сумм по градациям A_1 и A_2 (I, третий и шестой столбцы)	$\Sigma T^2_{АИ}=21^2+24^2+20^2+19^2+25^2+24^2+13^2+15^2=3373$
$\Sigma T^2_{ВИ}$	сумма квадратов индивидуальных сумм по градациям B_1 и B_2 (II, третий и шестой столбцы)	$\Sigma T^2_{ВИ}=26^2+27^2+20^2+16^2+20^2+21^2+13^2+18^2=3395$
n	количество испытуемых	$n=4$*1
a	количество градаций фактора А	$a=2$
b	количество градаций фактора В	$b=2$
N	общее количество индивидуальных значений	$N=16$
$(\Sigma x_i)^2$	квадрат общей суммы всех значений	$(\Sigma x_i)^2=161^2=25921$
$(\Sigma x_i)^2/N$	константа, которая вычитается из всех SS	$(\Sigma x_i)^2/N=161^2/16=1620,06$
Σx_i^2	сумма квадратов индивидуальных значений	$\Sigma x_i^2=11^2+13^2+12^2+9^2+10^2+11^2+8^2+10^2+15^2+14^2+8^2+7^2+10^2+10^2+5^2+8^2=1723$

Теперь при расчетах будем лишь подставлять уже подсчитанные значения тех или иных величин. В случае, если какой-то из шагов в алгоритме расчетов будет не вполне ясен, можно вернуться к Табл. 8.8 и восстановить процедуры расчетов, или к Табл. 8.6 и Табл. 8.7, для того, чтобы вспомнить, почему мы подставляем в формулу ту или иную конкретную величину.

На самом деле в эксперименте участвовало 20 человек. В дисперсионный комплекс случайным образом отобраны 4 из них в целях упрощения расчетов. Результаты дисперсионного анализа по такой

"усеченной" выборке совпадают с данными обработки всей выборки с помощью критерия χ^2_r.

Таблица 8.9

Последовательность операций в двухфакторном дисперсионном анализе для связанных выборок

	Операции	Формулы расчетов	Расчет по экспериментальным данным
1	Подсчитать SS_A	$SS_A=\Sigma T^2_A/_{n\cdot b}-(\Sigma x_i)^2/_N$	$SS_A=12985/(4\cdot 2)-1620,06=3,06$
2	Подсчитать SS_B	$SS_B=\Sigma T^2_B/_{a\cdot a}-(\Sigma x_i)^2/_N$	$SS_B=13105/(4\cdot 2)-1620,06=18,06$
3	Подсчитать $SS_{\text{И}}$	$SS_{\text{И}}=\Sigma T^2_{\text{И}}/_{a\cdot b}-(\Sigma x_i)^2/_N$	$SS_{\text{И}}=6665/(2\cdot 2)-1620,06=46,19$
4	Подсчитать SS_{AB}	$SS_{AB}=\Sigma T^2_{AB}/_n-(\Sigma x_i)^2/_N-SS_A-SS_B$	$SS_{AB}=6571/4-1620.06-3.06-$ $-18.06=1,56$
5	Подсчитать $SS_{A\text{И}}$	$SS_{A\text{И}}=\Sigma T^2_{A\text{И}}/b-(\Sigma x_i)^2/_N-SS_A-SS_{\text{И}}$	$SS_{A\text{И}}=3373/2-1620.06-3.06-$ $-46.19=17,19$
6	Подсчитать $SS_{B\text{И}}$	$SS_{B\text{И}}=\Sigma T^2_{B\text{И}}/a-(\Sigma x_i)^2/_N-SS_B-SS_{\text{И}}$	$SS_{B\text{И}}=3395/2-1620.06-18.06-$ $-46.19=13,19$
7	Подсчитать $SS_{\text{общ}}$	$SS_{\text{общ}}=\Sigma x^2_i-(\Sigma x_i)^2/_N$	$SS_{\text{общ}}=1723-1620.06=102,94$
8	Подсчитать $SS_{AB\text{И}}$	$SS_{AB\text{И}}=SS_{\text{общ}}-SS_A-SS_B-SS_{\text{И}}-$ $-SS_{AB}-SS_{A\text{И}}-SS_{B\text{И}}$	$SS_{AB\text{И}}=102,94-3,06-18,06-46,19-$ $-1,56-17,19-13,19=3,69$
9	Подсчитать число степеней свободы	$df_A=a-1$ $df_B=b-1$ $df_{\text{И}}=n-1$ $df_{AB}=df_A\cdot df_B$ $df_{A\text{И}}=df_A\cdot df_{\text{И}}$ $df_{B\text{И}}=df_B\cdot df_{\text{И}}$ $df_{AB\text{И}}=df_A\cdot df_B\cdot df_{\text{И}}$ $df_{\text{общ}}=N-1$	$df_A=2-1=1$ $df_B=2-1=1$ $df_{\text{И}}=4-1=3$ $df_{AB}=1\cdot 1=1$ $df_{A\text{И}}=1\cdot 3=3$ $df_{B\text{И}}=1\cdot 3=3$ $df_{AB\text{И}}=1\cdot 1\cdot 3=3$ $df_{\text{общ}}=16-1=15$
10	Разделить каждую SS на соответствующее число степеней свободы	$MS_A=SS_A/df_A$ $MS_B=SS_B/df_B$ $MS_{\text{И}}=SS_{\text{И}}/df_{\text{И}}$ $MS_{AB}=SS_{AB}/df_{AB}$ $MS_{A\text{И}}=SS_{A\text{И}}/df_{A\text{И}}$ $MS_{B\text{И}}=SS_{B\text{И}}/df_{B\text{И}}$ $MS_{AB\text{И}}=SS_{AB\text{И}}/df_{AB\text{И}}$	$MS_A=3,06/1=3,06$ $MS_B=18,06/1=18,06$ $MS_{\text{И}}=46,19/3=15,40$ $MS_{AB}=1,56/1=1,56$ $MS_{A\text{И}}=17,19/3=5,73$ $MS_{B\text{И}}=13,19/3=4,40$ $MS_{AB\text{И}}=3,69/3=1,23$
11	Подсчитать значения F и определить им df_1 по числителю и df_2 по знаменателю	$F_A=MS_A/MS_{A\text{И}}$ $F_B=MS_B/MS_{B\text{И}}$ $F_{\text{И}}=MS_{\text{И}}/MS_{AB\text{И}}$ $F_{AB}=MS_{AB}/MS_{AB\text{И}}$	$F_{A(1,3)}=3,06/5,73=0,53$ $F_{B(1,3)}=18,06/4,40=4,10$ $F_{\text{И}(3,3)}=15,40/1,23=12,52$ $F_{AB(1,3)}=1,56/1,23=1,27$
12	Определить критические значения F по Табл.XVII Приложения 1	На точке пересечения df_1 и df_2	$F_{\text{кр}(1,3)}=\begin{cases}10,13 & (p\le 0,05)\\ 34,12 & (p\le 0,01)\end{cases}$ $F_{\text{кр}(3,3)}=\begin{cases}9,28 & (p\le 0,05)\\ 29,46 & (p\le 0,01)\end{cases}$
13	Сопоставить эмпирические значения с критическими	При $F_{\text{эмп}}<F_{\text{кр}}$ H_0 принимается При $F_{\text{эмп}}\ge F_{\text{кр}}$ H_0 отклоняется	$F_A<F_{\text{кр}}\rightarrow H_0$ принимается $F_B<F_{\text{кр}}\rightarrow H_0$ принимается $F_{\text{И}}>F_{\text{кр}}\rightarrow H_0$ отклоняется ($p<0,05$) $F_{AB}<F_{\text{кр}}\rightarrow H_0$ принимается

Мы видим, что влияние факторов А и В, как каждого в отдельности, так и в их взаимодействии, незначимо. В то же время фактор индивидуальных различий между испытуемыми ($F_{\text{И}}$) оказался значимым ($p<0,05$). Мы видим из формы приведенного алгоритма, что этот индивидуальный источник вариативности с самого начала учитывается практически как третий фактор вариативности признака. Критерий F для факторов А и В вычисляется как отношение вариативности между градациями факторов к вариативности между испытуемыми в этих градациях.

На Рис. 8.3 индивидуальные изменения величин длительности физического волевого усилия представлены графически.

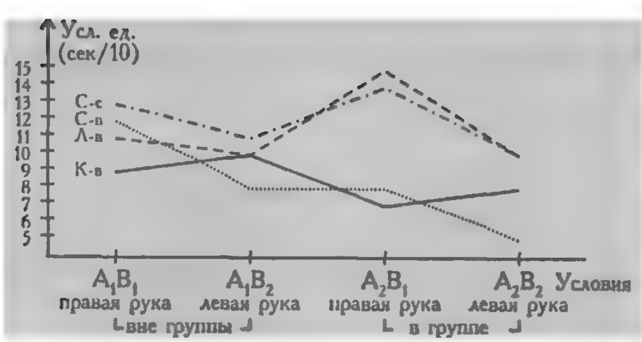


Рис. 8.3. Индивидуальные изменения длительности физического волевого усилия по четырем испытуемым

Как видно из Рис. 8.3, у одного испытуемого выше показатели по левой руке, у трех других - по правой. При измерении вне группы индивидуальные кривые ближе друг к другу, при измерениях в группе они расходятся. Можно было бы говорить об увеличении разброса индивидуальных значений при измерении длительности физического волевого усилия в группе, в атмосфере соревнования. Однако, несмотря на название, дисперсионный анализ выявляет влияние фактора не на рассеивание индивидуальных значений, а на среднюю их величину. Влияние же фактора на рассеивание признака можно уловить с помощью других критериев, в том числе непараметрических (Суходольский Г.В., 1972, с.341).

И все же представим полученный результат в принятой форме изменения средних значений по градациям факторов (Рис. 8.4).

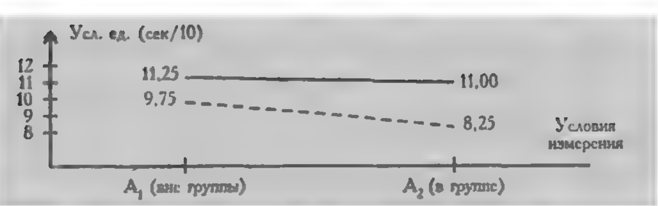


Рис. 8.4. Изменения средних величин длительности физического волевого усилия при переходе от индивидуальных замеров к групповым (правая рука - сплошная линия, левая рука - пунктирная линия)

Если исследователя интересует в большей степени второй вопрос данной задачи, связанный с проверкой предположения о том, что правая рука более "социальна", то он может представить данные в иной группировке (Рис. 8.5).

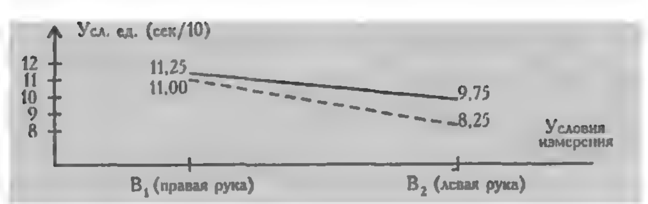


Рис. 8.5. Изменения средних величин длительности физического волевого усилия при переходе от правой руки к левой (сплошная линия - измерения вне группы, пунктирная линия - измерения в группе)

Мы видим, что во втором, групповом, замере снижаются показатели и по правой, и по левой руке, но все же правая рука "держится" почти на уровне первого замера, в то время как левая рука в большей степени "сдается" под влиянием усталости в группе, чем вне группы. Можно было бы подтвердить предположение о большей "социальности правой руки, большая стабильность которой, возможно, отражает стремление поддержать "лицо" в ситуации соревнования в группе, но выявленные тенденции, как мы убедились, незначимы.

Ограничения двухфакторного дисперсионного анализа для связанных выборок

Все ограничения такие же, как и в модели для несвязанных выборок, с одним уточнением. Все испытуемые должны пройти все сочетания градаций двух факторов. Этим достигается равномерность комплекса.

Итак, мы убедились, что двухфакторный дисперсионный анализ действительно позволяет нам оценить влияние двух факторов в их взаимодействии. Мы показали, что влияние одного фактора может оказаться различным при разных уровнях другого фактора, иногда различным вплоть до противоположности. Так, в примере о влиянии скорости предъявления слов и их длины на объем воспроизведения мы убедились в том, что фактор скорости при предъявлении коротких слов повышает результаты, а при предъявлении длинных слов - снижает результаты испытуемых.

Дисперсионный анализ позволяет также доказать, что влияние индивидуальных различий может оказаться сильнее экспериментальных или иных факторов, как это было продемонстрировано в последнем из примеров.

Более сложные схемы дисперсионного анализа позволяют анализировать совокупное действие трех, четырех и более факторов и получить еще более глубокие результаты.

ГЛАВА 9 РЕШЕНИЯ ЗАДАЧ С КОММЕНТАРИЯМИ

9.1. Рекомендации по решению задач

Лучше сначала попробовать решить задачу самостоятельно, выбрав критерий по алгоритму, приведенному в соответствующей главе.

Проверить правильность своего решения можно по ответам в настоящей главе.

Независимо от того, совпадает ли ваш ответ с приведенным в настоящей главе или нет, рекомендуется внимательно прочитать предлагаемое решение задачи. Дело в том, что в процессе анализа реальных исследовательских задач становится возможным проникнуть в те тонкости и дополнительные варианты использования статистических методов, которые в общем описании остаются "за кадром" рассмотрения.

Кроме того, способы интерпретации задач и тем более, интерпретации результатов также полнее раскрываются в описании решений, чем в формализованных изложениях процедур обработки.

9.2. Решения задач Главы 2

Решение задачи 1

Сопоставляются 2 выборки испытуемых. Следовательно, мы выбираем один из двух критериев: Q Роэенбаума или U Манна-Уитни.

Поскольку n_1, n_2<11, критерий Q не может быть использован (см. Алгоритм 7). Будем использовать критерий U Манна-Уитни. Если же он окажется бессильным выявить достоверные различия между группами, обратимся к угловому преобразованию Фишера - φ*.

Гипотезы лучше сформулировать после подсчета ранговых сумм. Предполагается, что в группе протагонистов показатели сокращения дистанции с оппонентами должны быть выше, чем в группе суфлеров, которые действовали лишь рационально, не вживаясь в роль оппонента. Однако лучше вначале определить, в какой из групп показатели не теоретически, а реально выше.

Будем действовать по алгоритму. Проранжируем все значения так, как если бы они принадлежали к одной общей выборке, а затем построим таблицу, в которой будут представлены индивидуальные значения и их ранги отдельно по двум группам (Табл. 9.1).

Таблица 9.1

Подсчет ранговых сумм по показателю сокращения психологической дистанции в группах протагонистов и суфлеров

Группа 1: протагонисты (n_1=7)			Группа 2: суфлеры (n_2=7)	
Показатель		Ранг	Показатель	Ранг
	75	14		
	50	13		
	30	11	30	11
	30 -	11		
	25	8,5	25	8,5
	20	6,5	20	6,5
			15	5
	10	3	10	3
			10	3
			5	1
Суммы	240	67	115	38
Средние	34,29		16,43	

Мы видим, что теоретические ожидания подтверждаются: в группе суфлеров ранговая сумма меньше.

Проверим, совпадает ли общая сумма рангов с расчетной величиной:

$\sum R_i$ = 67+38=105

$$\sum R_i = \frac{(n_1 + n_2 + 1) \cdot (n_1 + n_2)}{2} = \frac{(7 + 7 + 1) \cdot (7 + 7)}{2} = 105$$

Суммы совпадают. Мы можем перейти к формулированию гипотез.

H_0: Группа протагонистов (реальных исполнителей роли петербуржцев) не превосходит группы суфлеров по показателю сокращения психологической дистанции с оппонентами.

H_1: Группа протагонистов превосходит группу суфлеров по показатели сокращения психологической дистанции с оппонентами.

Определяем эмпирическое значение U:

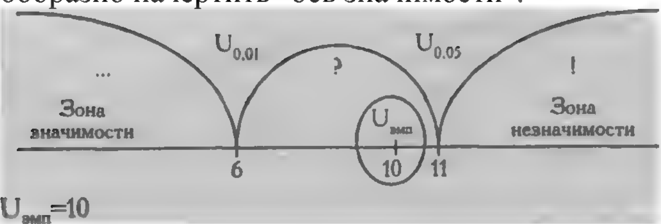
$$U = n_1 \cdot n_2 + \frac{n_x \cdot (n_x + 1)}{2} - T_x$$

$$U = 7 \cdot 7 + \frac{7 \cdot (7 + 1)}{2} - 67 = 49 + 28 - 67 = 10$$

Поскольку в данном случае $n_1 = n_2$, нам нет необходимости на всякий случай подсчитывать значение U для второй ранговой суммы. Определим I критические значение U по Табл. II Приложения 1 для $n_1=7$, $n_2=7$:

$$U_{кр} = \begin{cases} 11\,(p \le 0{,}05) \\ 6\,(p \le 0{,}01) \end{cases}$$

Критерий U - один из трех критериев, в которых меньшее значение свидетельствует о больших различиях. Для того, чтобы понять, достоверный ли мы получили результат, целесообразно начертить "ось значимости".



Это значение уже не попадает в "зону незначимости", но еще не попадает в "зону значимости". Но мы помним, что нас может удовлетворить и результат, соответствующий низшему порогу значимости: $p \le 0{,}05$.

$U_{эмп} < U_{кр}$ $(p<0{,}05)$

Ответ: H_0 отклоняется. Группа протагонистов превосходит группу суфлеров по показателю сокращения психологической дистанции с оппонентами ($p<0{,}05$).

Эти данные могли бы использоваться как еще одно подтверждение идеи Дж. Л. Морено о том, что принятие на себя роли оппонента способствует сближению с ним, если бы мы были уверены, что, во-первых, на роль протагонистов не вызвались участники изначально более расположенные к сближению с оппонентами, и что, во-вторых, испытуемые имели в виду одну и ту же дистанцию, когда определяли у себя процент ее сокращения. Впрочем, второе из этих ограничений распространяется и на большинство других шкал самооценки: мы не можем быть полностью уверены, что испытуемые оценивают у себя одно и то же качество или признак, как бы тщательно мы его ни определяли.

Данная задача является также примером сопоставления сдвигов в двух независимых выборках (см. параграф 3.1, Табл. 3.1).

Решение задачи 2

Поскольку в обеих выборках n_1, $n_2>11$ и диапазоны разброса значений в двух выборках не совпадают между собой, мы можем воспользоваться самым простым критерием для сопоставления двух выборок -критерием Q Розенбаума. Объемы выборок различаются менее чем на 10 человек, так что ограничение о примерном равенстве выборок также не препятствует нам.

Данные в Табл. 2.10 уже упорядочены по возрастанию признака. Первым, более высоким, рядом является ряд значений в мужской выборке.

Средняя величина тоже выше в выборке мужчин.

Сформулируем гипотезы.

H_0: При обращении в службу знакомств мужчинам приходится преодолевать не более интенсивное внутреннее сопротивление, чем женщинам.

H_1: При обращении в службу знакомств мужчинам приходится преодолевать более интенсивное внутреннее сопротивление, чем женщинам.

Сопоставим ряды значений для определения S_1 и S_2.

В Табл. 9.2 отмечены два интересующих нас значения: максимальное значение 2-го ряда (max 2) и минимальное значение 1-го ряда (min 1).

Определим S_1, как количество значений 1-го ряда, которые превышают максимальное значение 2-го ряда: S_1=5.

Определяем S_2, как количество значений 2-го ряда, которые меньше минимального значения 1-го ряда: S_2=5.

Вычисляем эмпирическое значение Q как суммы S_1 и S_2:

$Q=S_1+S_2=5+5=10$

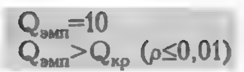
По Табл. I Приложения 1 определяем критические значения Q при n_1=17, n_2=23:

$$Q_{кр} = \begin{cases} 7\ (p \le 0{,}05) \\ 9\ (p \le 0{,}01) \end{cases}$$

Таблица 9.2 Расчет критерия Q при сопоставлении мужской (n_1=17) и женской (n_2=23) выборок по показателю интенсивности внутреннего сопротивления при обращении в службу знакомств

Группа 1 – мужчины (n_1=17)	Группа 2 - женщины (n_2=23)
S_1 81	
80	
73	
72	
72	
	max 2 70
69	
69	
	66
	66
65	
65	
	63
	63
62	61
60	60
54	54
54	
	47
43	43
	41
	40
	39
	38
	38
	35
30	30
	27
26	
26 min 1	
	25
	23
	17 S_2
	10

	9
Суммы 1001	965
Средние 58,89	41,96



Ответ: H_0 отвергается. Принимается H_1. При обращении в службу знакомств мужчинам из исследованной выборки пришлось преодолеть более мощное внутреннее сопротивление, чем женщинам.

Решение задачи 3

Поскольку мы сопоставляем 4 группы испытуемых, нам нужно выбирать между критерием тенденций S Джонкира и критерием Н Крускала-Уоллиса. В таких случаях мы должны сначала проверить, есть ли возможность применить первый из этих критериев, S, поскольку он позволяет не только выявить изменения, но и подтвердить направление этих изменений. В данном случае количество групп (*c*) меньше 6, количество испытуемых в каждой группе (*n*) меньше 10, при этом все группы численно равны. Следовательно, с формальной точки зрения критерий тенденций S применим. Вместе с тем, как мы можем определить по Табл. 2.11, показатели по фактору N при переходе от группы к группе изменяются *не* однонаправленно: сначала они возрастают, но в последней, четвертой, возрастной группе снижаются. На самом деле перед нами скорее не прямолинейная, а криволинейная зависимость (Рис. 9.1).



Рис. 9.1. Соотношение диапазонов значении и средних величин в четырех возрастных группах испытуемых по фактору N 16-факторного личностного опросника Р. Б. Кеттел-ла; для каждого диапазона указаны минимальное и максимальное значение в "сырых' баллах

Мы можем изменить последовательность расположения групп, упорядочив их по нарастанию значений фактора N, для чего придется поменять местами 4-ю и 3-ю группу.

Сформулируем гипотезы.

H_0: Тенденция возрастания значений по фактору N при переходе от группы к группе в последовательности 1-2-4-3 является случайной.

H_1: Тенденция возрастания значений по фактору N при переходе от группы к группе в последовательности 1-2-4-3 не является случайной.

Далее будем действовать по алгоритму 6 (Табл. 9.3).

Таблица 9.3

Расчет критерия S при сопоставлении разных возрастных групп по фактору N из ^6-факторного личностного опросника Р. Б. Кеттелла

№№ испыту-емых	Группа 1: 26-31 год (n_1=7)		Группа 2: 32-37 лет (n_2=7)		Группа 3 (ранее 4): 46-52 года (n_3=7)		Группа 4 (ранее 3): 38-42 года (n_4=7)	
	Индивиду-альные значения	Количество более высоких значениий справа	Индивиду-альные значения	Количестве более высоких значений справа	Индивиду-альные значения	Количество более высоких значений справа	Индивиду-альные значения	Количество более высоких значений справа
1	2	(21)	7	(14)	9	(5)	8	(0)
2	5	(21)	8	(13)	9	(5)	9	(0)
3	7	(20)	9	(Ю)	10	(4)	10	(0)
4	8	(18)	11	(7)	И	(4)	12	(0)
5	10	(12)	12	(5)	12	(3)	14	(0)
6	10	(12)	12	(5)	13	(3)	14	(0)
7	12	(5)	12	(5)	14	(1)	16	(0)
Суммы	54	(109)	71	(59)	78	(25)	83	(0)
Средние	7,71		10,14		11,11		11,86	

Определим величину А, которая является суммой всех чисел в скобках. Для этого просуммируем все суммы чисел в скобках по столбцам:

А=109+59+25=193

Теперь определим величину В по формуле:

$$B = \frac{c \cdot (c-1)}{2} \cdot n^2 = \frac{4 \cdot (4-1)}{2} \cdot 7^2 = 6 \cdot 49 = 294$$

Определяем эмпирическое значение S:

$S_{эмп}=2 \cdot A - B = 2 \cdot 193 - 294 = 92$

По Табл. IV Приложения 1 определяем критические значения для данного количества групп (c=4) и данного количества испытуемых в каждой группе (n=7):

$$S_{кр} = \begin{cases} 82\ (\rho \le 0{,}05) \\ 115\ (\rho \le 0{,}01) \end{cases}$$

$S_{эмп}=92$

$S_{эмп}>S_{кр}.\ (\rho \le 0{,}05)$

Ответ: H_0 отклоняется. Тенденция возрастания значений по фактору N не является случайной. Фактор N, отражающий житейскую искушенность и проницательность, имеет тенденцию возрастать при переходе от первой группы ко второй, а затем к четвертой; самые высокие значения приходятся на третью возрастную группу (от 38 до 42 лет).

Можем ли мы трактовать полученный результат в том смысле, что в период от 26 до 42 лет житейская искушенность и проницательность повышается, а 46-52 - снижается?

Нет, возрастные изменения признака может по-настоящему подтвердить только лонгитюдинальное многолетнее исследование одних и тех же испытуемых. В данном же случае мы выявили различия между возрастными группами по методу возрастных срезов, поэтому их можно объяснить, например, тем, что последняя возрастная группа (46-52 года) вообще является носителем иных ценностей и иных способов взаимодействия между людьми, при которых прямота, безыскусность и простодушие предпочтительнее изысканности, изощренности и хитрости.

Однако, учитывая малый объем выборки и низкий уровень значимости выявленной тенденции (p<0,05), такие выводы было бы делать слишком смело. Это лишь гипотеза, нуждающаяся в дальнейшей проверке.

Характерно, что применение критерия Н Крускала-Уоллиса дает в решении этой задачи незначимый результат.

Применение критерия Н Крускала-Уоллиса для решения задачи 3

Вначале сформулируем гипотезы.

H_0: Четыре возрастные группы испытуемых-руководителей промышленного

предприятия не различаются по уровню фактора N из 16PF[1].

H_1: Четыре возрастные группы испытуемых-руководителей промышленного предприятия различаются по уровню фактора N из 16PF.

В Табл. 9.4 реализованы первые шаги алгоритма в подсчете критерия Н.

[1] 16PF - принятое в иностранной н отечественной литературе сокращение для обозначения 16-факторного личностного опросника Р.Б. Кеттелла.

Таблица 9.4

Подсчет ранговых сумм по четырем возрастным группам испытуемых по фактору N из 16PF (N=28)

	Группа 1: 26-31 год (n_1=7)		Группа 2: 32-37 лет (n_2=7)		Группа 3: 38-42 года (n_3=7)		Группа 4: 46-52 года (n_4=7)	
	Индивиду-альные значения	Ранги	Индивиду-альные значения	Ранги	Индивиду-альные значения	Ранги	Индивиду-альные значения	Ранги
	2	1	7	3,5	8	6	9	9,5
	5	2	8	6	9	9,5	9	9,5
	7	3,5	9	9,5	10	13,5	10	13,5
	8	6	11	16,5	12	20,5	11	16,5
	10	13,5	12	20,5	14	26	12	20,5
	10	13,5	12	20,5	14	26	13	24
	12	20,5	12	20,5	16	28	14	26
Суммы *T*		60		97		129.5		119,5

Проверим, совпадает ли общая сумма рангов с расчетной величиной:

$$\sum R_i = \frac{N \cdot (N+1)}{2} = \frac{28 \cdot (28+1)}{2} = 406$$

$$\sum R_i = 60+97+129{,}5+119{,}5=406$$

Суммы равны, мы можем переходить к расчету эмпирического значения Н. Все расчеты будем выполнять с точностью до сотых долей единицы.

$$H = \left(\frac{12}{N \cdot (N+1)} \cdot \sum \frac{T_i^2}{n}\right) - 3 \cdot (N+1)$$

$$H = \left[\frac{12}{28 \cdot (28+1)} \cdot \left(\frac{60^2}{7} + \frac{97^2}{7} + \frac{129.5^2}{7} + \frac{119{,}5^2}{7}\right)\right] - 3 \cdot (28+1) = 6{,}15$$

Поскольку сопоставляется 4 группы испытуемых, а не 3, мы не можем воспользоваться специальной таблицей для критерия Н и должны обратиться к Табл. IX Приложения 1 для определения критических значений критерия χ^2_r. Для этого определим количество степеней свободы для данного количества групп (c=4):

$\nu = c - 1 = 4 - 1 = 3$

$$\chi^2_{кр} = \begin{cases} 7.815 \ (\rho \le 0{,}05) \\ 11.345 \ (\rho \le 0{,}01) \end{cases}$$

$$H_{эмп} = 6{,}15$$

$H_{эмп} < \chi^2_{кр.}$

Ответ: H_0принимается. Четыре возрастные группы руководителей промышленного предприятия не различаются по уровню фактора N 16-факторного личностного опросника Р.Б. Кеттелла.

Итак, мы смогли убедиться в том, что критерий Н оказывается менее мощным, чем критерий S Джонкира. Это еще один аргумент в пользу того, чтобы во всех тех случаях, когда это возможно, при сопоставлении 3 и более выборок отдавать предпочтение критерию тенденций S.

9.3. Решения задач Главы 3

Решение задачи 4

Оценки отношения к наказаниям определены для 3-х условий, и вопрос задачи требует проверки достоверности тенденции в оценках. Целесообразнее всего было бы

использовать критерий тенденций L Пейджа, но количество испытуемых п=16, а критические значения критерия L определены только для $n \leq 12$. Используем вначале критерий Фридмана, а затем все же попробуем использовать критерий L, разделив выборку на 2 части.

Решение задачи с использованием критерия χ^2_r Фридмана

Сформулируем гипотезы:

H_0: Испытуемые примерно в одинаковой степени оправдывают (признают возможными) телесные наказания, которые их ребенок может получить от них самих, от бабушки и от воспитательницы (или учительницы).

H_1: Испытуемые в разной степени оправдывают телесные наказания; которые их ребенок может получить от них самих, от бабушки и от воспитательницы (или учительницы).

Проранжируем оценки каждого испытуемого по трем условиям. Ранжирование производится по строкам, при этом меньшая оценка получает меньший ранг, большая оценка - наибольший ранг (Табл. 9.5).

Таблица 9.5

Оценки допустимости телесных наказаний со стороны разных людей и их ранги (n=16)

Испытуемые	Условие 1: "Я сам"		Условие 2: "Бабушка"		Условие 3: "Учительница"	
	Оценка	Ранг	Оценка	Ранг	Оценка	Ранг
1	4	3	2	2	1	1
2	1	2	1	2	1	2
3	5	3	4	1.5	4	1.5
4	4	3	3	2	2	1
5	3	2.5	3	2.5	2	1
6	4	2	5	3	1	1
7	3	2.5	3	2.5	1	1
8	5	2,5	5	2.5	3	1
9	6	3	5	2	3	1
10	2	2	2	2	2	2
11	6	3	3	2	2	1
12	5	3	3	1	4	2
13	7	3	5	2	4	1
14	5	2,5	5	2,5	2	1
15	5	2,5	5	2,5	4	1
16	6	2.5	6	2.5	4	1
Суммы	71	42	60	34,5	40	19,5
Средние	4.44		3,75		2,50	

Как видно из Табл. 9.5, суммы рангов по каждому условию составляют: 42; 34,5; 19,5, что в сумме равняется 96.

Расчетная сумма рангов:

$$\sum R_i = n \cdot \frac{c \cdot (c+1)}{2} = 16 \cdot \frac{3 \cdot (3+1)}{2} = 96$$

Реально полученная и расчетная суммы совпадают, мы можем переходить к дальнейшим расчетам.

Определим эмпирическое значение χ^2_r:

$$\chi^2_{r\,эмп} = \left[\frac{12}{n \cdot c \cdot (c+1)} \cdot \sum T_j^2\right] - 3 \cdot n \cdot (c+1)$$

$$\chi^2_{r\,эмп} = \left[\frac{12}{16 \cdot 3 \cdot (3+1)} \cdot \left(42^2 + 34{,}5^2 + 19{,}5^2\right)\right] - 3 \cdot 16 \cdot (3+1) = 16{,}4$$

В нашем случае количество условий c=3, однако $n>9$, поэтому мы не можем воспользоваться таблицами, специально рассчитанными для критерия χ^2_r. Нам придется сопоставлять полученное эмпирическое значение с критическими значениями критерия χ^2_r. Число степеней свободы определяем по формуле:

$v=c-1=3-1=2$.

По Табл. IX Приложения 1 определяем критические значения для v=2:

$$\chi^2_{кр} = \begin{cases} 5{,}991\ (\rho \le 0{,}05) \\ 9{,}210\ (\rho \le 0{,}01) \end{cases}$$

$\chi^2_{r\ эмп}=16{,}4$

$\chi^2_{r\ эмп}>\chi^2_{r\ кр}$ $(\rho\le0{,}01)$

Ответ: H_0 отклоняется. Принимается H_1. Испытуемые в разной степени оправдывают телесные наказания, которые их ребенок может получить от них самих, от бабушки и от воспитательницы (учительницы).

Решение с использованием критерия тенденций L Пейджа

Из Табл. 9.5 видно, что испытуемые, похоже, склонны более снисходительно относиться к тем наказаниям, которые они сами дают детям *(T_1=42),* несколько менее снисходительно они относятся к бабушкиным наказаниям (T_2=34,5), и еще менее снисходительно - к наказаниям со стороны воспитательницы или учительницы, хотя бы и "за дело (T_3=19,5). Метод Пейджа требует, чтобы мы расположили условия в порядке возрастания ранговых сумм: условия 1, 2 и 3 становятся, соответственно, условиями 3, 2 и 1, как показано в Табл. 9.6.

Имеющиеся таблицы критических значений критерия L рассчитаны только для небольших выборок ($n\le12$). В исследованной выборке n=16. Попробуем обойти это ограничение следующими двумя способами:

1) Разделим выборку пополам и рассчитаем отдельно для каждой подгруппы из 8 человек эмпирическое значение критерия L. Если в обоих случаях будут выявлены достоверные тенденции изменения оценок, мы сможем распространить этот вывод на выборку в целом.

2) Напишем на карточках условные номера всех 16 испытуемых, перемешаем карточки, перевернув их лицевой стороной вниз, а затем случайным образом отберем 12 испытуемых и рассчитаем для них эмпирическое значение критерия L. Этот метод применяется в дисперсионном анализе для уравновешивания комплексов (см. Главы 7 и 8).

Мы можем применить в данном случае и сам дисперсионный анализ, но ограничимся пока этими двумя способами.

Таблица 9.6

Оценки допустимости телесных наказаний и их ранги в упорядоченной для критерия L последовательности (n_1=8; n_2=8)

Испытуемые	Условие 1 (бывшее 3): "Учительница"		Условие 2 (бывшее 2): "Бабушка"		Условие 3 (бывшее 1): "Я сам"	
	Оценка	Ранг	Оценка	Ранг	Оценка	Ранг
1	1	1	2	2	4	3
2	1	2	1	2	1	2
3	4	1,5	4	1,5	5	3
4	2	1	3	2	4	3
5	2	1	3	2,5	3	2,5
6	1	1	5	3	4	2
7	1	1	3	2,5	3	2,5
8	3	1	5	2,5	5	2,5
Суммы	15	9,5	26	18	29	20,5
Средние	1,875		3,25		3,63	
9	3	1	5	2	6	3
10	2	2	2	2	2	2
И	2	1	3	2	6	3
12	4	2	3	1	5	3
13	4	1	5	2	7	3
14	2	1	5	2,5	5	2,5
15	4	1	5	2,5	5	2,5

16	4	1	6	2,5	6	2,5
Суммы	25	10	34	16,5	42	21,5
Средние	3,125		4,25		5,25	

Сформулируем гипотезы.

H_0: Повышение оценок допустимости телесных наказаний от первого условия к третьему случайно.

H_1: Повышение оценок допустимости телесных наказаний от первого условия к третьему не случайно.

Определим L_1 и L_2 для двух половин нашей выборки по формуле:

$L_1=\sum(T_j \cdot j)=(9,5\cdot1)+(18\cdot2)+(20,5\cdot3)=9,5+36+61,5=107$

$L_2=\sum(T_j \cdot j)=(10\cdot1)+(16,5\cdot2)+(21,5\cdot3)=10+33+64,5=107,5$

По Табл. VIII Приложения 1 определяем критические значения L для $n=8$, c=3:

$$\chi^2_{кр}=\begin{cases}5,991\ (\rho\le0,05)\\9,210\ (\rho\le0,01)\end{cases}$$

$\chi^2_{r\ эмп}=16,4$

$\chi^2_{r\ эмп}>\chi^2_{r\ кр}\ (\rho\le0,01)$

$$L_{кр}=\begin{cases}104\ (\rho\le0,05)\\106\ (\rho\le0,01)\\109\ (\rho\le0,001)\end{cases}$$

Построим "ось значимости"



Мы видим, что для обеих половин выборки $L_{эмп}>L_{кр}$, что позволяет нам отвергнуть нулевую гипотезу ($p\le0,01$).

Теперь используем второй способ сокращения выборки.

Случайным образом отобраны 12 испытуемых из 16: 1, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16. Все расчеты для этой усеченной выборки представлены в Табл. 9.7.

Таблица 9.7

Расчет критерия L по оценкам допустимости телесных наказаний для усеченной выборки испытуемых (n=12)

Испытуемые		Условие 1: "Учительница"		Условие 2: "Бабушка"		Условие 3: "Я сам"	
		Оценка	Ранг	Оценка	Ранг	Оценка	Ранг
1	№1	1	1	2	2	4	3
2	№3	4	1,5	4	1.5	5	3
3	№4	2	1	3	2	4	3
4	№5	2	1	3	2,5	3	2,5
5	№6	1	1	5	3	4	2
6	№7	1	1	3	2,5	3	2,5
7	№8	3	1	5	2,5	5	2,5
8	№9	3	1	5	2	6	3
9	№10	2	2	2	2	2	2
10	№12	4	2	3	1	5	3
11	№14	2	1	5	2,5	5	2,5
12	№16	4	1	6	2,5	6	2,5
Суммы		29	14,5	46	26	52	31,5

Средние	2,42		3,83		4,33	

$L_{эмп}$=(14,5·1)+(26·2)+(31,5·3)=14,5+52+94,5=161

По Табл. VIII Приложения 1 определяем критические значения L для n=12, c=3:

$$L_{кр} = \begin{cases} 153 \ (\rho \le 0{,}05) \\ 156 \ (\rho \le 0{,}01) \\ 160 \ (\rho \le 0{,}001) \end{cases}$$

$L_{эмп} > L_{кр}$ (ρ<0,001)

Ответ: H_0 отклоняется. Принимается H_1. Повышение оценок от первого условия к третьему неслучайно (p<0,001). Испытуемые менее всего склонны соглашаться на то, чтобы воспитательница или учительница применяла телесное наказание по отношению к их ребенку, более склонны соглашаться с тем, чтобы это делала бабушка и еще более склонны позволять это делать себе.

Но, конечно, когда мы говорим о меньшей или большей склонности, то ориентируемся на эмпирически установленный диапазон значений и средние величины, которые "на глаз" не так уж сильно различаются и составляют, соответственно: 2,50; 3,75; 4,44 по 7-балльной шкале.

Решение задачи 5

Вопрос 1: Ощущаются ли участниками значимые сдвиги в уровне владения каждым из трех навыков после тренинга?

Поскольку данные представлены по одной экспериментальной выборке и было совершено 2 замера, мы должны выбирать между критерием знаков G и критерием T Вилкоксона (см. Алгоритм 12). Сдвиги могут быть определены количественно, но они варьируют в достаточно узком диапазоне - от –2 до +4. Учитывая это, применим последовательно оба критерия.

Использование критерия знаков для решения задачи 5.

Преобразуя данные Табл. 3.9, составим таблицу сдвигов, которая будет полезна при использовании обоих критериев. Для этого из значения, полученного во 2-м замере, вычтем значение, полученное данным испытуемым по данной шкале в 1-м замере.

Таблица 7.7

Последовательность операций в однофакторной модели дисперсионного анализа для связанных выборок

Операция	Формула расчета	Расчет по экспериментальным данным
1 Подсчитать $SS_{\text{факт}}$	$SS_{\text{факт}}={}^1/_n\cdot\Sigma T^2_c-$ $-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_{\text{факт}}={}^1/_5\cdot(51^2+1244^2+47^2)-{}^1/_{15}\cdot1342^2=$ $={}^1/_5\cdot1552346-{}^1/_{15}\cdot1800964=190405$
2 Подсчитать $SS_{\text{исп}}$	$SS_{\text{исп}}={}^1/_c\cdot\Sigma T^2_и-$ $-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_{\text{исп}}={}^1/_3\cdot(247^2+631^2+100^2+181^2+183^2)-$ $-{}^1/_{15}\cdot1342^2={}^1/_5\cdot535420-{}^1/_{15}\cdot1800964=$ $=58409$
3 Подсчитать $SS_{\text{общ}}$	$SS_{\text{общ}}=\Sigma x^2_i-$ $-{}^1/_N\cdot(\Sigma x_i)^2$	$SS_{\text{общ}}=5^2+7^2+2^2+2^2+35^2+235^2+604^2+$ $+93^2+171^2+141^2+7^2+20^2+5^2+8^2+7^2-$ $-{}^1/_{15}\cdot1800964=479706-120064,26=$ $=359642$
4 Подсчитать $SS_{\text{сл}}$	$SS_{\text{сл}}=SS_{\text{общ}}-SS_{\text{факт}}-$ $-SS_{\text{исп}}$	$SS_{\text{сл}}=359642-190405-58409=110828$
5 Подсчитать число степеней свободы	$df_{\text{факт}}=c-1$ $df_{\text{исп}}=n-1$ $df_{\text{общ}}=N-1$ $df_{\text{сл}}=df_{\text{общ}}-df_{\text{факт}}-df_{\text{исп}}$	$df_{\text{факт}}=3-1=2$ $df_{\text{исп}}=5-1=4$ $df_{\text{общ}}=15-1=14$ $df_{\text{сл}}=14-2-4=8$
6 Разделить каждую SS на число степеней свободы	$MS_{\text{факт}}=SS_{\text{факт}}/df_{\text{факт}}$ $MS_{\text{исп}}=SS_{\text{исп}}/df_{\text{исп}}$ $MS_{\text{сл}}=SS_{\text{сл}}/df_{\text{сл}}$	$MS_{\text{факт}}=190405/2=95202,5$ $MS_{\text{исп}}=58409/4=14602,2$ $MS_{\text{сл}}=110827/8=13853,4$
7 Подсчитать значения F и определить им df_1 по числителю и df_2 по знаменателю	$F_{\text{факт}}=MS_{\text{факт}}/MS_{\text{сл}}$ $F_{\text{исп}}=MS_{\text{исп}}/MS_{\text{сл}}$	$F_{\text{факт}(2,8)}=95202,5/13853,4=6,872$ $F_{\text{исп}(4,8)}=14602,2/13853,4=1,054$
8 Определить критические значения F по Табл. XVII Приложения 1	для $df_1=2$ и $df_2=8$ для $df_1=4$ и $df_2=8$	$F_{\text{кр}\,(2,8)}=\begin{cases}4,46\ (p\le0,05)\\8.65\ (p\le0,01)\end{cases}$ $F_{\text{кр}\,(2,8)}=\begin{cases}3,84\ (p\le0,05)\\7,01\ (p\le0,01)\end{cases}$
9 Сопоставить эмпирические значения F с критическими	При $F_{\text{эмп}}<F_{\text{кр}}$ H_0 принимается При $F_{\text{эмп}}\ge F_{\text{кр}}$ H_0 отклоняется	$F_{\text{факт}}>F_{\text{кр}}\rightarrow H_{0(А)}$ отклоняется $F_{\text{исп}}<F_{\text{кр}}\rightarrow H_{0(Б)}$ принимается

Вывод: $H_{0(А)}$ отклоняется. Различия в объеме воспроизведения слов в разных условиях являются более выраженными, чем различия, обусловленные случайными причинами ($p<0,05$).

$H_{0(Б)}$ принимается: Индивидуальные различия между испытуемыми являются не более выраженными, чем различия, обусловленные Случайными причинами.

Однако, судя по Рис. 7.3, мы не можем утверждать, что срабатывает фактор длины анаграммы. Более значимыми оказываются качественные, а не количественные различия между анаграммами. Как мы уже имели возможность убедиться (см. параграфы 3.4 и 3.5), непараметрический L - критерий Пейджа подтверждает тенденцию увеличения Вщдивидуальных показателей при переходе от анаграммы КРУА к анаграмме ИНААМШ, а затем к анаграмме АЛСТЬ ($p<0,01$). Значимые различия были получены и с помощью критерия Фридмана χ^2_r ($p=0,0085$).

Итак, непараметрические критерии позволяют нам констатировать Вэолее высокий уровень значимости различий между условиями! В Зачем же тогда использовать достаточно сложный дисперсион-Вшй анализ? Для того, чтобы подобрать существенные факторы, кото-Врые могут стать основой для формирования двух-, трех- и более фак-Вгорных дисперсионных комплексов, позволяющих оценить не только В^лияние каждого из факторов в отдельности, но и их взаимодействие.

Таблица 9.8

Сдвиги в уровне владения коммуникативными навыками "после"-"до" ($n=12$)

Код имени испытуемого		Сдвиги			Ранги абсолютных величин сдвигов		
		Активное слушание	Снижение напряжения	Аргументация	Активное слушание	Снижение напряжения	Аргументация
1	Ис.	1	1	2	3	3	6,5

2	Я.	2	3	1	8	10,5	3
3	Ин.	4	3	1	11,5	10,5	3
4	Р.	2	1	0	8	3	–
5	К.	**–2**	1	1	**8**	3	3
6	Н.	2	2	3	8	7	8,5
7	Ен.	4	3	3	11,5	10,5	8,5
8	Ле.	**–1**	2	2	**3**	7	6,5
9	Ли.	1	1	0	3	3	–
10	Т.	2	1	1	8	3	3
11	Ет.	**–1**	**–2**	0	**3**	**7**	–
12	Б.	1	3	1	3	10,5	3
Количество нетипичных сдвигов		3	1	0	Суммы рангов нетипичных сдвигов		
		12	12	9	12	12	0
Всего сдвигов		12	12	9	12	12	9[2]

Из Табл. 9.8 мы видим, что положительных сдвигов по всем шкалам больше.

Сформулируем гипотезы.

H_0: Преобладание типичного (положительного) сдвига в самооценках уровня владения коммуникативными навыками является случайным.

H_1: Преобладание типичного (положительного) сдвига в самооценках уровня владения коммуникативными навыками не является случайным.

[2] Как мы помним, "нулевые" сдвиги исключаются из рассмотрения, а количество сопоставляемых пар значений уменьшается на величину, соответствующую количеству таких сдвигов; в данном случае для шкалы "Аргументация" п—9.

[3] Суммы рангов для шкал "Активное слушание" и "Снижение напряжения" составляют 78, что совпадает с расчетной суммой

$$\sum R_i = \frac{N\cdot(N+1)}{2} = \frac{12\cdot(12+1)}{2} = 78$$

Общая сумма рангов для шкалы "Аргументация" составляет 45, что также совпадает с расчетной суммой

$$\sum R_i = \frac{9\cdot(9+1)}{2} = 45$$

Проверим гипотезы, определив критические значения критерия знаков по Табл. V Приложения 1.

Для шкалы "Активное слушание", n=12:

$$G_{кр} = \begin{cases} 2\,(p \le 0{,}05) \\ 1\,(p \le 0{,}01) \end{cases}$$

$G_{эмп}=3$

$G_{эмп}>G_{кр}$

Ответ: H_1 принимается.

Для шкалы "Снижение напряжения", n=12:

$$G_{кр} = \begin{cases} 2\,(p \le 0{,}05) \\ 1\,(p \le 0{,}01) \end{cases}$$

$G_{эмп}=1$

$G_{эмп}\le G_{кр}$

Ответ: H_0 отвергается. Принимается H_1. Преобладание положительного сдвига в самооценке уровня владения навыком снижения напряжения не является случайным ($p\le 0{,}01$).

Для шкалы "Аргументация", n=9:

$$C_{кр} = \begin{cases} 1\ (p \le 0{,}05) \\ 0\ (p \le 0{,}01) \end{cases}$$

$G_{эмп}=0$

$G_{эмп} \le G_{кр}$

Ответ: H_0 отвергается. Принимается H_1. Преобладание положительного сдвига в самооценке уровня владения навыком аргументации не является случайным ($p \le 0{,}01$).

Итак, преобладание положительного сдвига по навыку активного слушания является случайным, а по навыкам снижения эмоционального напряжения и аргументации не является случайным ($p \le 0{,}01$).

Использование критерия Т Вилкоксона для решения задачи 5

Сформулируем гипотезы.

H_0: Интенсивность положительных сдвигов не превосходит интенсивности отрицательных сдвигов.

H_1: Интенсивность положительных сдвигов превосходит интенсивность отрицательных сдвигов.

В Табл. 9.8 нами уже просуммированы ранги "редких", в данном случае, отрицательных, сдвигов. Сопоставляем эти значения с максимальными значениями Т, при которых различия еще могут считаться достоверными (Табл. VI Приложения 1).

Для шкалы "Активное слушание", $n=12$:

$$T_{кр} = \begin{cases} 17\ (p \le 0{,}05) \\ 9\ (p \le 0{,}01) \end{cases}$$

$T_{эмп}=14$

$T_{эмп} < T_{кр}$

Ответ: H_0 отклоняется. Преобладание положительных сдвигов по навыкам активного слушания неслучайно ($p<0{,}05$). Для шкалы "Снижение напряжения", $n=12$:

$$T_{кр} = \begin{cases} 17\ (p \le 0{,}05) \\ 9\ (p \le 0{,}01) \end{cases}$$

$T_{эмп}=7$

$T_{эмп} < T_{кр}$

Ответ: H_0 отклоняется. Принимается H_1. Преобладание положительных сдвигов по навыку снижения напряжения не является случайным ($p<0{,}01$).

Для шкалы "Аргументация", $n=9$:

$$T_{кр} = \begin{cases} 8\ (p \le 0{,}05) \\ 3\ (p \le 0{,}01) \end{cases}$$

$T_{эмп}=0$

$T_{эмп} < T_{кр}$

Ответ: H_0 отвергается. Принимается H_1. Преобладание положительных сдвигов по навыкам аргументации неслучайно ($p<0{,}01$).

Итак, участники ощущают значимые положительные сдвиги по всем трем группам коммуникативных навыков.

В данном случае критерий Т доказал свою большую мощность по сравнению с критерием знаков. Он подтвердил ранее установленные различия на высоком уровне значимости ($p<0{,}01$) и позволил выявить их для шкалы "Активное слушание" ($p<0{,}05$).

Однако мы не можем интерпретировать полученные результаты в терминах эффективности тренинга по меньшей мере по двум причинам:

1) у нас не было контрольной группы, у которой измерялись бы те же показатели с тем же интервалом времени;

2) показатели самооценки после тренинга могли отражать желание испытуемых косвенно поблагодарить тренера за его работу.

Несмотря на это, все-таки есть смысл ответить на второй вопрос задачи, проверив, различаются ли между собой величины сдвигов по трем разным шкалам. Со всеми возможными поправками на индивидуальные тенденции к завышению или занижению

самооценок, различия в сдвигах все же отражают относительную эффективность тренинговых воздействий по трем направлениям.

Вопрос 2: Произошли ли по трем видам навыков разные сдвиги или эти сдвиги для разных навыков примерно одинаковы?

Величины сдвигов получены по трем разным шкалам для одной и той же выборки испытуемых. Для того, чтобы определить, различаются ли величины сдвигов, полученных по трем шкалам, применимы критрии χ^2_r Фридмана и L Пейджа.

Таблица 9.9

Сдвиги в оценках уровня развития коммуникативных навыков и их ранги (n=12)

Код имени	Оценка	Признак 1: Активное слушание		Признак 2: Снижение эмоционального напряжения		Признак 3: Аргументация	
		Оценка	Ранг	Оценка	Ранг	Оценка	Ран
1	Ис.	1	1,5	1	1,5	2	3
2	Я.	2	2	3	3	1	1
3	Ин.	4	3	3	2	1	1
4	Р.	2	3	1	2	0	1
5	К.	**–2**[4]	**1**	**1**	**2.5**	**1,5**	**2,5**
6	Н.	2	1,5	2	1,5	3	3
7	Ен.	4	3	3	1,5	3	1,5
8	Ле.	**–1**	**1**	**2**	**2,5**	**2**	**2,5**
9	Ли.	1	2,5	1	2,5	0	1
10	Т.	2	3	1	1,5	1	1,5
11	Ет.	**–1**	**2**	**–2**	**1**	**0**	**3**
12	Б	1	1,5	3	3	1	1,5
	Суммы	15	25 (21)	19	24,5 (18,5)	15	22,5 (1
	Средние	1,25		1,58		1,25	

[4] Отрицательную величину считаем меньшей величиной и приписываем ей, соответственно, меньший ранг. Может получиться так, что большую величину ранга -третий ранг - получит значение 0, как это имеет место у испытуемого Ет. (№11). В каком-то смысле при двух отрицательных сдвигах третий нулевой сдвиг является положительным, но это можно и оспаривать. Поэтому целесообразно рассчитать значение L отдельно для всех испытуемых и для тех испытуемых, у кого нет отрицательных сдвигов (п=9). Соответствующие суммы приведены в скобках.

Проранжируем сдвиги по трем шкалам для каждого испытуемого (Табл. 9.9). Ранжирование, как мы помним, производится по строкам.

Поскольку количество замеров c=3, т. е. меньше 6, а количество испытуемых гг=12, мы можем остановить выбор на критерии тенденций L Пейджа. Такая возможность благоприятна, так как критерий L по мощности превосходит критерий χ^2_r (см., например, задачу 3 и ее решение).

Проверим соответствие сумм рангов расчетным суммам. Сумма рангов по всей выборке составляет 25+24,5+22,5=72. Расчетная сумма:

$$\sum R_i = n \cdot \frac{c \cdot (c+1)}{2} = 12 \cdot \frac{3 \cdot (3+1)}{2} = 72$$

Сумма рангов по усеченной выборке (n=9) составляет 21+18,5+14,5=54. Расчетная сумма:

$$\sum R_i = n \cdot \frac{c \cdot (c+1)}{2} = 9 \cdot \frac{3 \cdot (3+1)}{2} = 54$$

В обоих случаях суммы рангов совпадают с расчетными, мы можем перейти к дальнейшим действиям.

Сформулируем гипотезы, ориентируясь на значения ранговых сумм;

H_0: Тенденция к меньшему сдвигу по шкале "Аргументация", промежуточному

сдвигу по шкале "Снижение напряжения" и большему сдвигу по шкале "Активное слушание" является случайной.

H_1: Тенденция к меньшему сдвигу по шкале "Аргументация", промежуточному сдвигу по шкале "Снижение напряжения" и большему сдвигу по шкале "Активное слушание" не является случайной.

Определим эмпирические значения критерия L по всей выборке в целом:

$L_{эмп} = \sum(T_j \cdot j) = (22{,}5 \cdot 1) + (24{,}5 \cdot 2) + (25 \cdot 3) = 22{,}5 + 49 + 75 = 146{,}5$

По Табл. VIII Приложения 1 определяем критические значения L для n=12, c=3:

$$L_{кр} = \begin{cases} 153\ (\rho \le 0{,}05) \\ 156\ (\rho \le 0{,}01) \\ 160\ (\rho \le 0{,}001) \end{cases}$$

$L_{эмп} = 146{,}5$

$L_{эмп} < L_{кр}$

H_0 принимается.

Определим эмпирическое значение критерия L для усеченной выборки:

$L_{эмп} = (14{,}5 \cdot 1) + (18{,}5 \cdot 2) + (21 \cdot 3) = 14{,}5 + 37 + 63 = 114{,}5$

Определяем по Табл.VIII Приложения 1 критические значения L при n =9:

$$L_{кр} = \begin{cases} 116\ (\rho \le 0{,}05) \\ 119\ (\rho \le 0{,}01) \\ 121\ (\rho \le 0{,}001) \end{cases}$$

$L_{эмп} = 114{,}5$

$L_{эмп} < L_{кр}$

H_0 принимается.

Ответ: H_0 принимается и для полной, и для усеченной выборки. Тенденция к меньшему сдвигу по шкале "Аргументация", промежуточному сдвигу по шкале "Снижение напряжения" и наибольшему сдвигу по шкале "Активное слушание" является случайной.

Итак, общий вывод таков: сдвиги в показателях по трем видам коммуникативных навыков достоверны, но указать, в каком из видов навыков участники ощущают больший сдвиг, а в каком - меньший, на основании этих данных невозможно.

Вопрос 3: Уменьшается ли расхождение между "идеальным" и реальным уровнями владения навыками после тренинга?

Сокращение расхождения между индивидуальным идеалом и самооценкой - один из главных показателей эффективности психотерапевтического воздействия (Rogers C.,1961,p.236; Роджерс К.,1995,с.292). Сближение самооценки реального Я и идеального Я происходит в большинстве случаев аа счет повышения реальной самооценки, но может снизиться и уровень идеальных требований к себе благодаря переключению на более реалистичные и менее "наказующие" цели.

Итак, мы проверяем, оказал ли тренинг психотерапевтическое воздействие на участников. Как правило, испытуемые не предполагают, что у них измеряется не абсолютный уровень самооценки или "идеала" и, даже, не расхождение между ними, а расхождение между расхождениями, сдвиг в величине этого расхождения после тренинга. Можно предположить, что этот показатель более объективно отражает происходящие изменения. По крайней мере, он в меньшей степени подвергнут влиянию фактора социальной желательности.

Поскольку мы сопоставляем 2 разных представляемых или умозрительных условия измерения на одной и той же выборке испытуемых и по одному и тому же набору показателей, применимы критерии знаков и T Вилкоксона.

Поскольку расхождения варьируют в достаточно широком диапазоне - от 3 до 5, целесообразнее использовать критерий T Вилкоксона.

В Табл. 9.10 по каждой шкале представлены 4 показателя: расхождение между

идеальным и реальным уровнями до тренинга, после тренинга, разность между расхождениями "после" и "до" и ранги этих разностей (сдвигов).

Таблица 9.10

Сдвиг в величинах расхождения между "идеалом" и реальным уровнем развития коммуникативных навыков

Код имени частника		Активное слушание				Снижение напряжения				Аргументация			
		До	после	Сдвиг (после - до)	Ранг сдвига	до	после	Сдвиг Ранг (после сдвига — до)		до	после	Сдвиг (после - до)	Ранг сдвига
1	Ис.	3	3	0	—	3	4	–	**3.5**	3	2	-1	3,5
2	Я.	2	2	0	—	2	2	0	-	1	2	**1**	**3,5**
3	Ин.	2	2	0	—	2	1	-1	3,5	3	2	-1	3,5
4	Р.	2	1	-1	2	1	2	**1**	**3,5**	2	2	0	–
5	К.	3	6	**3**	**4,5**	5	5	0	-	4	5	**1**	**3,5**
6	Н.	2	1	-1	2	3	2	-1	3,5	3	2	-1	3,5
7	Ен.	5	1	-4	6	5	2	-3	7	4	2	-2	7
8	Ле.	3	3	0	-	3	3	0	-	4	4	0	–
9	Ли.	2	1	-1	2	4	3	-1	3.5	4	4	0	—
10	Т.	3	3	0	-	3	3	0	-	3	4	**1**	**3,5**
11	Ет.	2	5	**3**	**4,5**	4	5	**1**	**3,5**	6	6	0	—
12	Б.	2	2	0	_	7	2	-5	8	3	3	0	–
Всего сдвигов				6				8				7	
Типичный сдвиг		Отрицательный				Отрицательный				Отрицательный			
Сумма рангов нетипичных сдвигов					9				10,5				10.5

В Табл. 9.10 выделены величины нетипичных, более редко встречающихся, сдвигов, и ранги их абсолютных значений. Мы видим, что большинство сдвигов - это нулевые или отрицательные сдвиги. Это означает, что расхождение между идеалом и самооценкой чаще уменьшается или остается на прежнем уровне, чем увеличивается. Однако нас сейчас интересует именно уменьшение расхождения между идеальным и реальным Я, а поэтому все нулевые сдвиги придется исключить из рассмотрения.

Сформулируем гипотезы.

H_0: Сближение идеального и реального уровней навыков после тренинга не является преобладающей тенденцией.

H_1: Сближение идеального и реального уровней навыков после тренинга является преобладающей тенденцией.

Сближение выражается в отрицательном, типичном, сдвиге расхождения между идеальным и реальным уровнями.

По Табл. V Приложения 1 определяем критические значения критерия Т и сопоставляем их с эмпирическими значениями. По шкале "Активное слушание"", *n*=6:

$$T_{кр} = \begin{cases} 2\ (p \le 0{,}05) \\ 0\ (p \le 0{,}01) \end{cases}$$

$T_{эмп} = 9$

$T_{эмп} > T_{кр}$
Нд принимается.
По шкале "Снижение напряжения", *n*=8:
$$T_{кр} = \begin{cases} 5\,(p \le 0{,}05) \\ 1\,(p \le 0{,}01) \end{cases}$$
$T_{эмп} = 10{,}5$
$T_{эмп} > T_{кр}$
Но принимается.
По шкале "Аргументация", п=7:
$$T_{кр} = \begin{cases} 3\,(p \le 0{,}05) \\ 0\,(p \le 0{,}01) \end{cases}$$
$T_{эмп} = 10{,}5$
$T_{эмп} > T_{кр}$
Но принимается.

Ответ: Т - критерий Вилкоксона не позволяет отвергнуть нулевую гипотезу. Уменьшение расхождения между идеальным и реальным уровнями навыков не является доминирующей тенденцией.

Исследователь может утешать себя тем, что в процессе тренинга участники ощутили новые горизонты развития... Действительно, произошли достоверные положительные сдвиги не только в оценке *реального* уровня владения коммуникативными навыками (см. выше), но и достоверные положительные сдвиги в оценке *идеального* уровня. Кроме того, в исследованиях К. Роджерса речь идет не о самооценке уровня владения коммуникативными навыками, **а** о более глубоких аспектах личностной самооценки в методе Q - сортировки. Учитывая малый объем выборки, полученный результат можно считать лишь предварительным.

9.4. Решения задач Главы 4

Решение задачи 6

Вопрос 1: Можно ли утверждать, что разные картины методики Хекхаузена обладают разной побудительной силой в отношении мотивов: а) "надежда на успех"; б) "боязнь неудачи"?

Для того, чтобы ответить на этот вопрос, необходимо сопоставить распределение реакций "надежда на успех" и реакций "боязнь неудачи" с равномерным распределением. Тем самым мы проверим, равномерно ли распределяются реакции "надежды на успех" по шести картинам и равномерно ли распределяются реакции "боязни неудачи" по шести картинам.

Количество наблюдений достаточно велико, чтобы мы могли использовать любой из классических критериев - χ^2 или λ. Однако, как мы помним, картины в данном исследовании предъявлялись разным испытуемым в разных последовательностях, следовательно, мы не можем говорить об однонаправленном изменении признака в какую-либо одну сторону: все разряды (картины) следуют друг за другом в случайном порядке. Это является веским основанием для применения критерия χ^2 и отказа от критерия λ.

Рассмотрим оба аспекта поставленного вопроса последовательно.

А) Равномерно ли распределяются реакции "надежды на успех" по шести картинам методики Хекхаузена?

H_0: Распределение реакций "надежды на успех" не отличается от равномерного распределения.

H_1: Распределение реакций "надежды на успех" отличается от равномерного распределения.

Рассчитаем теоретические частоты для равномерного распределения по формуле:
$f_{теор} = n/k$,
где *n* - количество наблюдений,
k - количество разрядов.

В данном случае количество наблюдений - это количество реакций "надежды на успех" у 113 испытуемых. Таких реакций зарегистрировано 580, следовательно, n =580. Количество разрядов - это количество стимульных картин, следовательно, $k=6$. Определяем $f_{теор}$:

$f_{теор}={}^{n}/_{k}={}^{580}/_{6}=96,7$

Количество степеней свободы V определяем по формуле:

$v= k -1=6-1=5$

Итак, поправка на непрерывность не нужна, мы можем производить все расчеты по общему алгоритму. Они представлены в Табл.9.11.

Таблица 9.11

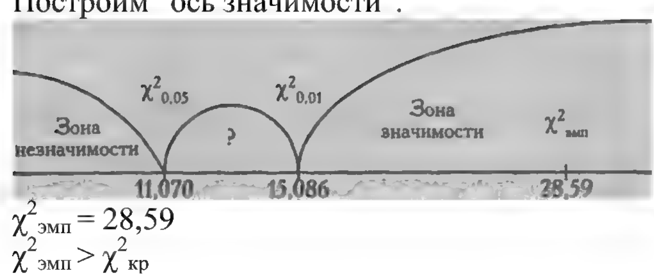
Расчет критерия χ^2 при сопоставлении распределения реакций "надежды на успех" по 6 картинам с равномерным распределением

Разряды-картины методики		Эмпирические частоты реакций "надежды на успех" $f_э$	Теоретические частоты реакции "надежды на успех" $f_т$	$f_э- f_т$	$(f_э- f_т)^2$	$(f_э- f_т)^2/ f_т$
1	"Мастер измеряет деталь"	106	96,67	9,33	87,05	0,90
2	"Преподаватель и ученик"	102	96,67	5,33	28,41	0,29
3	"В цехе у машины"	108	96,67	11,33	128,37	1,33
4	"У двери директора"	50	96,67	-46,67	2178,09	22,53
5	"Человек в бюро"	99	96,67	2,33	5,43	0,06
6	"Улыбающийся юноша"	115	96,67	18,33	335,99	3,48
Суммы		580		0		28,59

По Табл. IX Приложения 1 определяем критические значения χ^2 для v=5:

$$\chi^2_{кр} = \begin{cases} 11,070 \ (\rho \le 0,05) \\ 15,086 \ (\rho \le 0,01) \end{cases}$$

Построим "ось значимости".



$\chi^2_{эмп} = 28,59$

$\chi^2_{эмп} > \chi^2_{кр}$

Ответ: H_0 отклоняется. Принимается H_1. Распределение реакций "надежды на успех" по шести картинам методики Хекхаузена отличается от равномерного распределения (p<0,01).

Б) Равномерно ли распределяются реакции "боязни неудачи" по шести картинам методики Хекхаузена?

H_0: Распределение реакций "боязни неудачи" не отличается от равномерного распределения.

H_1: Распределение реакций "боязни неудачи" отличается от равномерного распределения.

В данном случае количество наблюдений - это число реакций "боязни неудачи", следовательно, n=516; количество разрядов - это число стимульных картин, как и в предыдущем случае, следовательно, $k=6$. Определяем $f_{теор}$

$f_{теор} ={}^{516}/_{6}=86$

Количество степеней свободы $v=k—1=6—1=5$. Поправка на непрерывность здесь тоже, естественно, не нужна.

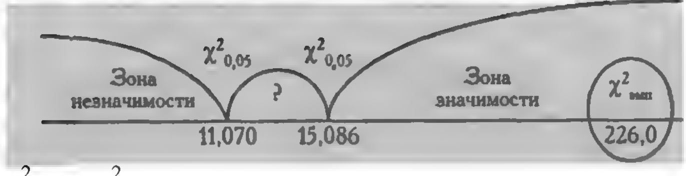
Все дальнейшие расчеты проделаем по алгоритму в таблице.

Таблица 9.12

Расчет критерия при сопоставлении распределения реакций "боязни неудачи" по 6 картинам с равномерным распределением

Разряды-картины методики		Эмпирические частоты реакций "боязни неудачи" $f_э$	Теоретические частоты реакции "боязни неудачи" $f_т$	$f_э- f_т$	$(f_э- f_т)^2$	$(f_э- f_т)^2/ f_т$
1	"Мастер измеряет деталь"	138	86	52	2704	31,44
2	"Преподаватель и ученик"	180	86	94	8836	102.74
3	"В цехе у машины"	34	86	-52	2704	31,44
4	"У двери директора"	87	86	1	1	0.01
5	"Человек в бюро"	57	86	-29	841	9.78
6	"Улыбающийся юноша"	20	86	-66	4356	50,65
Суммы		516	516	0	19442	226,06

Критические значения χ^2 при $v=5$ по Таблице IX Приложения 1 нам уже известны:



$\chi^2_{эмп} > \chi^2_{кр}$

Ответ: H_0 отклоняется. Принимается H_1. Распределение проявлений "боязни неудачи" по шести стимульным картинам отличается от равномерного распределения ($p<0,01$).

Итак, реакции "надежды на успех" и реакции "боязни неудачи" неравномерно проявляются в ответ на 6 стимульных картин. Однако это еще не означает, что эти картины являются неуравновешенными по направленности воздействия. Может оказаться так, по крайней мере теоретически, что одни и те же картины вызывают большинство реакций обоих типов, а другие картины почти не вызывают реакций или вызывают их достоверно меньше. В этом случае оба эмпирических распределения отличались бы от равномерного, но не различались бы между собой.

Проверим, различаются ли картины теперь уже не по количеству вызываемых реакций, а по их качеству, то есть вызывают ли одни картины скорее реакции "надежды на успех", а другие - реакции "боязни неудачи"

Вопрос 2: Можно ли считать стимульный набор методики Хекхаузена неуравновешенным по направленности воздействия?

Решим эту задачу двумя способами: а) путем сравнения распределения реакций "надежда на успех" с распределением реакций "боязнь неудачи" по 6-и картинам; б) путем сопоставления распределения реакций на каждую картину с равномерным распределением.

Выясним, совпадают ли распределения реакций по двум картинам. Для этого сформулируем гипотезы.

H_0: Распределения реакций "надежда на успех" и реакций "боязнь неудачи" не различаются между собой.

H_1: Распределения реакций "надежда на успех" и "боязнь неудачи" различаются между собой.

Для того, чтобы облегчить себе задачу подсчета теоретических частот, воспроизведем таблицу эмпирических частот и дополним ее.

Таблица 9.13

Эмпирические и теоретические частоты распределения реакций "надежда на успех" и "боязни неудачи"

Разряды - картины		Эмпирические частоты				Суммы	Теоретические частоты				Суммы
		Реакций "надежда на успех"		Реакций "боязнь неудачи"			Реакций "надежда на успех"		Реакций "боязнь неудачи"		
1	"Мастер измеряет деталь"	106	А	138	Б	244	129,1	А	114,9	Б	244
2	"Преподаватель и ученик"	102	В	180	Г	282	149,2	В	132,8	Г	282
3	"В цехе у маши-	108	Д	34	Е	142	75,1	Д	66,9	Е	142
4	"У двери директора"	50	Ж	87	З	137	72,5	Ж	64,5	З	137
5	"Человек в бюро"	99	И	57	К	156	82,6	И	73,4	К	156
6	"Улыбающийся юноша"	115	Л	20	М	135	71,4	Л	63,6	М	135
Суммы		580		516		1096	580		516		1096

Расчет теоретических частот осуществляется по известной нам формуле:

$$f_{теор} = \left(\text{Сумма частот по соответствующей строке}\right) \cdot \left(\frac{\text{Сумма частот по соответствующему столбцу}}{\text{Общее количество наблюдений}}\right)$$

Произведем расчеты.

$f_{А\ теор}=244 \cdot 580/1096=129,1$

$f_{Б\ теор}=244 \cdot 516/1096=114,9$

$f_{В\ теор}=282 \cdot 580/1096=149,2$

$f_{Г\ теор}=282 \cdot 516/1096=132,8$

$f_{Д\ теор}=142 \cdot 580/1096=75,1$

$f_{Е\ теор}=142 \cdot 516/1096=66,9$

$f_{Ж\ теор}=137 \cdot 580/1096=72,5$

$f_{З\ теор}=137 \cdot 516/1096=64,5$

$f_{И\ теор}=156 \cdot 580/1096=82,6$

$f_{К\ теор}=156 \cdot 516/1096=73,4$

$f_{Л\ теор}=135 \cdot 580/1096=71,4$

$f_{М\ теор}=135 \cdot 516/1096=63,6$

По Табл. 9.13 мы видим, что сумма всех теоретических частот равна общему количеству наблюдений, а попарные суммы теоретических частот по строкам равны суммам наблюдений по строкам.

Расчеты критерия χ^2 будем производить по известному алгоритму. Поправка на непрерывность не вносится, так как $v>1$:

$v=(r-1)(c-1)=(6-1)(2-1)=5$

Результаты всех операций по Алгоритму 13 представлены в Табл. 9.14.

Таблица 9.14

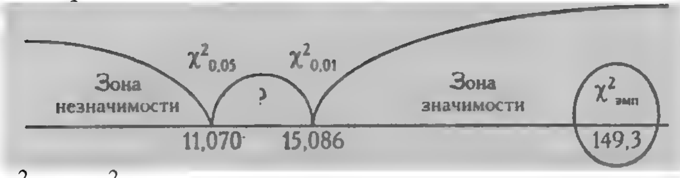
Расчет критерия χ^2 при сопоставлении эмпирических распределений реакций "надежды на успех" (НУ) и "боязни неудачи" (БН)

	Ячейки таблицы частот	Эмпирическая частота $f_э$	Теоретическая частота $f_т$	$f_э - f_т$	$(f_э - f_т)^2$	$(f_э - f_т)^2 / f_т$
1	А	106	129,1	-23,1	533,61	4,13
2	Б	138	114,9	23,1	533,61	4,64
3	В	102	149,2	-47,2	2227,84	14,93
4	Г	180	132,8	47,2	2227,84	16,78
5	Д	108	75,1	32,9	1082,41	14,41
6	Е	34	66,9	-32,9	1082,41	16,18
7	Ж	50	72,5	-22,5	506,25	6,98
8	З	87	64,5	22,5	506,25	7,85
9	И	99	82,6	16,4	268,96	3,26
10	К	57	73,4	-16,4	268,96	3,66
11	Л	115	71,4	43,6	1900,96	26,62
12	М	20	63,6	-43,6	1900,96	29,89
Суммы ,		1096	1096	0		149,33

Критические значения χ^2 при v=5 нам уже известны:

$$\chi^2_{кр} = \begin{cases} 11{,}070\ (\rho \le 0{,}05) \\ 15{,}086\ (\rho \le 0{,}01) \end{cases}$$

Построим "ось значимости".



$\chi^2_{эмп} > \chi^2_{кр}$

Ответ: H_0 отвергается. Принимается H_1. Распределения реакций "надежды на успех" и "боязни неудачи" различаются между собой.

Теперь выясним, совпадают ли распределения реакций по каждой картине. Сформулируем гипотезы.

H_0: Реакции двух видов в ответ на картину №1 (№2, №3 ... №6) распределяются равномерно.

H_1: Реакции двух видов в ответ на картину №1 (№2, №3 ... №6) распределяются неравномерно.

Реакции "надежды на успех" будем обозначать как НУ, реакции "боязни неудачи" - как БН.

Подсчитаем теоретические частоты для каждой из шести картин, по формуле:

$f_{теор} = n/k$,

где *n* общее количество реакций обоих направлений на данную картину; *k* - количество разрядов, в данном случае количество видов реакции (k =2).

$f_{1\ теор}$ =244/2=121;

$f_{2\ теор}$ =282/2=141;

$f_{3\ теор}$ =142/2=71;

$f_{4\ теор}$ =137/2=68,5

$f_{5\ теор}$ =156/2=78

$f_{6\ теор}$ =135/2=67,5

В данном случае число степеней свободы v=l:

$v=k—1=2—1=1$.

Следовательно, мы должны сделать во всех шести случаях поправку на непрерывность. Проведем расчеты отдельно для каждой картины (см. Табл. 9.15).

Таблица 9.15

Расчет критерия χ^2 при сопоставлении распределений реакций на каждую из шести картин с равномерным распределением

№	Вид реакции	Эмпирические частоты $f_э$	Теоретические частоты $f_т$	$(f_э-f_т)$	$(\|f_э-f_т\|-0,5)$	$(\|f_э-f_т\|-0,5)^2$	$\frac{(\|f_э-f_т\|-0,5)^2}{f_т}$
1	НУ	106	122	– 16,0	15,5	240,25	1,97
	БН	138	122	+ 16,0	15,5	240,25	1,97
	Суммы	244	244	0			3,94
2	НУ	102	141	– 39,0	38,5	1482,25	10,51
	БН	180	141	+ 39,0	38,5	1482,25	10,51
	Суммы	282	282	0			21,02
3	НУ	108	71	+ 37,0	36,5	1332,25	18,76
	БН	34	71	– 37,0	36,5	1332,25	18,76
	Суммы	142	142	0			37,52
4	НУ	50	68,5	–18,5	18,0	324,0	4,73
	БН	87	68,5	+ 18,5	18,0	324,0	4,73
	Суммы	137	137	0			9,46
5	НУ	99	78	+ 21,0	20,5	420,25	5,39
	БН	57	78	– 21,0	20.5	420,25	5,39
	Суммы	156	156	0			10,78
6	НУ	115	67,5	+ 47,5	47,0	2209,00	32,73
	БН	20	67,5	– 47,5	47,0	2209,00	32,73
	Суммы	135	135	0			65,46

Определим по Табл. IX Приложения 1 критические значения для ν=l:

$$\chi^2_{кр} = \begin{cases} 3{,}841\ (\rho \le 0{,}05) \\ 6{,}635\ (\rho \le 0{,}01) \end{cases}$$

$$\chi^2_{эмп(1)} > \chi^2_{кр}\ (\rho \le 0{,}05)$$

$$\chi^2_{эмп(2,3,4,5,6)} > \chi^2_{кр}\ (\rho \le 0{,}01)$$

Ответ: H_0 отклоняется для всех картин. H_1 принимается для картин 2, 3, 4, 5 и 6: реакции двух видов в ответ на эти картины распределяются неравномерно.

Если представить данные графически (Рис. 9.2), то легко можно видеть, что картины №6, №3 и №5 вызывают достоверно больше реакций "надежды на успех", а картины №2, №1 и №4 - достоверно больше реакций "боязни неудачи".

Стимульный набор методики X. Хекхаузена оказался неуравновешенным по направленности стимулирующего воздействия.

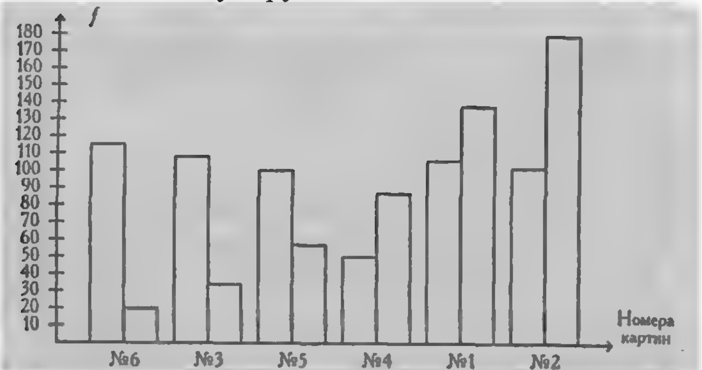


Рис. 9.2. Соотношения частот реакций "надежда на успех" (незаштрихованные столбн-ки) и реакций "боязнь неудачи" (заштрихованные столбики) по разным картинам методики Х.Хекхаузена

Вместе с тем, из Рис. 9.2 мы можем заметить, что если частоты реакций "боязни неудачи" достаточно монотонно возрастают при переходе от картины №6 к картине №3, а затем к №5, №4, №1 и №2, то частоты реакций "надежда на успех" по всем картинам, за исключением картины №4, оказываются примерно на одном уровне, в диапазоне от 99 до 115. Каждый исследователь сам для себя решает вопрос о том, что для него важнее - абсолютные показатели стимулирующего воздействия или их соотношения. Метод *у}* поможет ему решить задачи и первого, и второго типа.

Решение задачи 7

Вопрос 1: Можно ли утверждать, что распределение запретов не является равномерным?

Поскольку количество разрядов (запретов) $k>3$, и перечень из пяти запретов представляет собой номинативную шкалу, мы можем использовать только критерий χ^2.

Если бы участники тренинга называли разные запреты с одинаковой частотой, то каждый из пяти запретов встречался бы равновероятно с остальными.

Сформулируем гипотезы.

H_0: Распределение частот встречаемости пяти запретов не отличаетсяот равномерного распределения.

H_1: Распределение частот встречаемости пяти запретов отличается от равномерного распределения.

Определим $f_{теор}$ по формуле:

$f_{теор} = n/k$,

где n - общее количество наблюдений, в данном случае названных запретов (n =281); k - количество категорий запретов (k =5).

$f_{теор}$ =281/5=56,2

Определим число степеней свободы v:

$v = k$ -l=5-l=4.

Поправки на непрерывность делать не требуется. Все расчеты представим в таблице, строго следуя Алгоритму 13.

Таблица 9.16

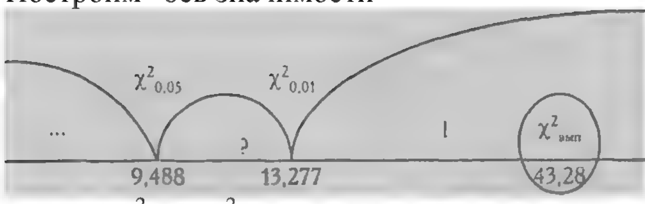
Расчет критерия χ^2 при сопоставлении эмпирического распределения частот встречаемости 5-и психологических запретов с равномерным распределением

Разряды - вид запрета	Эмпирическая частота $f_э$	Теоретическая частота $f_т$	$f_э - f_т$	$(f_э - f_т)^2$	$(f_э - f_т)^2/f_т$
1, Не давай психологических поглаживаний	44	56,2	-12,2	148,8	2,65
2. Не принимай...	45	56,2	-11,2	125,4	2,23
3. Не проси...	98	56,2	+41,8	1747,2	31,09
4. Не отказывайся...	58	56,2	+1,8 -	3,2	0,06
5. Не давай себе...	36	56,2	20,2	408,0	7,26
Суммы	281	281	0		43,29

Определим критические значения χ^2 по Таблице IX Приложения 1 для v=4:

$$\chi^2_{кр} = \begin{cases} 9{,}488 \ (\rho \leq 0{,}05) \\ 13{,}277 \ (\rho \leq 0{,}01) \end{cases}$$

Построим "ось значимости"



Ответ: $\chi^2_{эмп} > \chi^2_{кр}$ ($p \leq 0{,}01$)

H_0 отклоняется. Принимается H_1. Распределение частот встречаемости пяти психологических запретов отличается от равномерного распределения ($p<0{,}01$).

Вопрос 2: Можно ли утверждать, что запрет "Не проси" встречается достоверно чаще остальных?

Для того, чтобы ответить на этот вопрос, мы можем попробовать сопоставить запрет

"Не проси" последовательно со всеми остальными запретами, объединяя их попарно.

H_0: Распределение выборов между запретами "Не проси" и "Не давай" не отличается от равномерного распределения.

H_1: Распределение выборов между запретами "Не проси" и "Не давай" отличается от равномерного распределения.

Аналогичные гипотезы могут быть сформулированы для всех остальных пар запретов.

При сопоставлении двух запретов число разрядов *k=2,* следовательно, количество степеней свободы *v=k*—1=1. Это означает, что нам необходимо делать поправку на непрерывность.

Рассчитаем теоретические частоты для каждой из сопоставляемых пар запретов.

$f_{теор}=n/k,$

где *n* - сумма частот, приходящихся на данную пару запретов; *k* - количество сопоставляемых категорий запретов *(k*=2*)*.

Определим теоретические частоты для всех возможных пар запретов.

$f_{теор\ 1\text{-}2}$=(44+45)/2=44,5
$f_{теор\ 1\text{-}3}$=(44+98)/2=71
$f_{теор\ 1\text{-}4}$=(44+58)/2=51
$f_{теор\ 1\text{-}5}$=(44+36)/2=40
$f_{теор\ 2\text{-}3}$=(45+98)/2=71,5
$f_{теор\ 2\text{-}4}$=(45+58)/2=51,5
$f_{теор\ 2\text{-}5}$=(45+36)/2=40,5
$f_{теор\ 3\text{-}4}$=(98+58)/2=78
$f_{теор\ 3\text{-}5}$=(98+36)/2=67
$f_{теор\ 4\text{-}5}$=(58+36)/2=47

Теперь подсчитаем значения критерия χ^2 (Табл. 9.17).

Таблица 9.17. Расчет значений критерия при попарном сопоставлении частот запретов

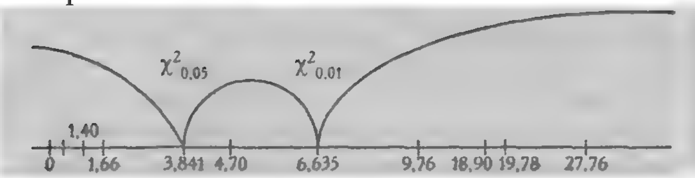
Сопоставляемые виды запретов		Эмпирические частоты $f_э$	Теоретические частоты $f_т$	$(f_э - f_т)$	$(\|f_э - f_т\| - 0{,}5)$	$(\|f_э - f_т\| - 0{,}5)^2$	$\frac{(\|f_э - f_т\| - 0{,}5)^2}{f_т}$
1	«Не давай»	44	44,5	-0,5	0	0	0
2	«Не принимай	45	44,5	+ 0,5	0	0	0
	Суммы	99	99,0	0			0
1	«Не давай»	44	71,0	-27,0	26,5	702,25	9,89
3	«Не проси»	98	71,0	+ 27,0	26,5	702,25	9,89
	Суммы	142	142,0	0			19,78
1	«Не давай»	44	51,0	-7,0	6,5	42,25	0,83
4	«Не отказывайся»	58	51,0	+ 7,0	6,5	42,25	0,83
	Суммы	102	102,0	0			1.66
1	«Не давай»	44	40,0	+ 4,0	3,5	12,25	0,31
5	«Не давай себе»	36	40,0	-4,0	3,5	12,25	0,31
	Суммы	80	80,0	0			0,62
2	«Не принимай»	45	71,5	-26,5	26,0	676,00	9,45
3	«Не проси»	98	71,5	+26,5	26,0	676,00	9,45
	Суммы	143	143,0	0			18,90
2	«Не принимай»	45	51,5	-6,5	6,0	36,00	0,70
4	«Не отказывайс	58	51,5	+ 6,5	6,0	36,00	0,70

	я» Суммы	103	103,0	0			1,40
2 5	«Не принимай»	45	40,5	+ 4,5	4,0	16,00	0,40
	«Не давай себе»	36	40,5	-4,5	4,0	16,00	0,40
	Суммы	81	81,0	0			0,80
3 4	«Не проси»	98	78,0	+ 20,0 -	19,5	380,25	4,88
	«Не отказывайся» Суммы	58 156	78,0 156,0	20,0 0	19,5	380,25	4,88 9,76
3 5	«Не проси»	98	67,0	+ 31,0 -	30,5	930,25	13,88
	«Не давай себе*	36	67,0	31,0	30,5	930,25	13,88
	Суммы	134	134,0	0			27,76
4 5	«Не отказывайся»	58	47,0	+ 11,0	10,5	110,25	2,35
	«Не давай себе»	36	47,0	-11,0	10,5	110,25	2,35
	Суммы	94	94,0	0			4,70

Определим критические значения χ^2 для v =1:

$$\chi^2_{кр} = \begin{cases} 3,841\ (p \le 0,05) \\ 6,635\ (p \le 0,01) \end{cases}$$

Построим "ось значимости".



Мы видим, что в некоторых случаях $\chi^2_{эмп} > \chi^2_{кр}$, а в некоторых - $\chi^2_{эмп} < \chi^2_{кр}$.

Мы можем суммировать полученные данные, построив матрицу, в которой какими-либо знаками будет отмечено, являются ли различия между данной парой запретов достоверными или недостоверными. Например, это могут быть указания на уровень значимости различий.

Запреты	1 запрет	2 запрет	3 запрет	4 запрет	5 запрет
1 запрет	—	—	p<0,01	—	—
2 запрет		—	p<0,01	—	—
3 запрет			—	p<0,01	p<0,01
4 запрет				—	p<0,05
5 запрет					—

Итак, выявлены достоверные различия в частоте встречаемости запрета 3 по сравнению со всеми остальными запретами (р<0,01 во всех четырех случаях) и запрета 4 по сравнению с запретом 5 (р<0,05).

Ответ: Hq отклоняется для пар запретов 1—3, 2—3, 3—4, 3—5 (p<0,01) и пары 4—5 (p<0,05). Запрет "Не проси психологических поглаживаний от других людей" встречается достоверно чаще, чем все остальные четыре запрета (p<0,01). Запрет "Не давай психологических поглаживаний самому себе" встречается реже, чем запрет "Не отказывайся от психологических поглаживаний, даже если они тебе не нравятся" (p<0,05). Обсуждение этих данных представлено в другой работе (Сидоренко Е. В., 1995, с. 65-67).

Решение задачи 8

Вопрос 1: Различаются ли распределения предпочтений, выявленные по каждому из четырех типов мужественности, между собой?

Для выявления различий между четырьмя распределениями лучше всего применить критерий χ^2. Критерий λ не применим по трем причинам: 1) n<50; 2) разряды представляют собой номинативную шкалу, так как при переходе от типа к типу

изменяется "качество", а не "количество" мужественности; 3) критерий λ позволяет сопоставлять только 2 распределения одновременно, а в нашу задачу входит одновременное сопоставление четырех распределений.

Сформулируем гипотезы.

H_0: Распределения предпочтений, выявленные по четырем типам мужественности, не различаются между собой.

H_1: Распределения предпочтений, выявленные по четырем типам мужественности, различаются между собой.

Рассчитаем теоретические частоты для каждой ячейки таблицы эмпирических частот (Табл. 9.18) по формуле:

$$f_{теор} = \left(\text{Сумма частот по соответствующей строке}\right) \cdot \left(\frac{\text{Сумма частот по соответствующему столбцу}}{\text{Общее количество наблюдений}}\right)$$

$f_{A\ теор}$ =31-31/124=7.75

$f_{Б\ теор}$ =31-31/124=7,75 и т. д.

Поскольку суммы по всем строкам и столбцам таблицы равны, теоретические частоты для всех 16-ти ячеек таблицы будут одинаковыми. Равенство же по строкам и столбцам объясняется тем, что каждая испытуемая совершала принужденный выбор, так что каждый из типов мужественности был выбран 31 раз (даже если он был "выбран на последнее место).

Эта задача напоминает шутливый литературный пример, в котором одна невеста совершала выбор из четырех женихов. В данном же случае у нас 31 испытуемая, и каждая совершает выбор из четырех типов мужественности, распределяя их по четырем позициям.

Определим количество степеней свободы V для четырех типов мужественности *(k)* и четырех позиций выбора (*c*):

v=(k-l)-(c-l)=(4-l) (4-l)=3·3=9

Все дальнейшие расчеты произведем в таблице по Алгоритму 13 без поправки на непрерывность, так как при v>1 она не требуется.

Таблица 9.18

Расчет критерия χ^2 при сопоставлении распределений четырех типов мужественности по четырем позициям (n=31)

Разряды-типы мужественности	Позиции выбора	Эмпирическая частота $f_э$	Теоретическая частота $f_т$	$f_э - f_т$	$(f_э - f_т)^2$	$(f_э - f_т)^2 / f_т$
1. Мифологически тип	1	2	7,75	-5,75	33,063	4,266
	2	6	7,75	-1,75	3,063	0,395
	3	4	7,75	-3,75	14,063	1,815
	4	19	7,75	+11,25	126,563	16,331
2. Национальный ТИП	1	19	7,75	+11,25	126,563	16,331
	2	4	7,75	-3,75	14,063	1,815
	3	7	7,75	-0,75	0,563	0,073
	4	1	7,75	-6,75	45,563	5,879
3, Современный ТИП	1	7	7,75	-0,75	0,563	0,073
	2	10	7,75	+2,25	5,063	0,653
	3	12	7,75	+4.25	18,063	2,331
	4	2	7,75	-5,75	33,063	4,266
4, Религиозный ТИП	1	3	7,75	-4,75	22,563	2,911
	2	11	7,75	+3,25	10,563	1,362
	3	8	7,75	+0,25	0,063	0,008
	4	9	7,75	+1,25	1,563	0,202
Суммы		124	124,0	0		58,711

По Табл. IX Приложения 1 определяем критические значения χ^2 при V=9:

$$\chi^2_{кр} = \begin{cases} 16{,}919\ (\rho \le 0{,}05) \\ 21{,}666\ (\rho \le 0{,}01) \end{cases}$$

$\chi^2_{эмп}=58{,}71$

$\chi^2_{эмп}>\chi^2_{кр}\ (\rho\le0{,}01)$

Ответ: H_0 отвергается. Принимается H_1. Распределения предпочтений по четырем типам мужественности различаются между собой.

Вопрос 2. Можно ли утверждать, что предпочтение отдается какому-то одному или двум типам мужественности? Наблюдается ли какая-либо групповая тенденция предпочтений?

В данном случае удобнее всего применить критерий χ^2_r Фридмана (см. Главу 3). Как мы помним, он позволяет выявить изменения в величине признака при переходе от одного условия к другому. По-видимому, еще более целесообразным было бы применить тест тенденций L Пейджа, но при n>12 это можно сделать только с помощью специальных ухищрений (см. Задачу 4 и ее решение).

Критерий χ^2_r позволяет определить, достоверным ли образом различаются суммы рангов, полученные по каждому из рассматриваемых условий, в данном случае - по каждому типу мужественности.

При этом ранги начисляются отдельно по каждому испытуемому, а суммируются - по каждому условию. В нашем случае нет необходимости что-то ранжировать, так как каждая испытуемая своими выборами фактически уже проранжировала четыре исследуемых типа мужественности. Суммы рангов по каждому типу мужественности можно подсчитать, умножая значение ранга на количество рангов с данным значением. Например, из Табл. 9.18 следует, что Мифологический тип 2 раза оказался в первой позиции. Значит, сумма рангов по 1-й позиции будет равна: 1·2=2. На второй позиции он оказался 6 раз, следовательно, сумма рангов по 2-й позиции равна: 2·6=12 и т. д. Произведем расчеты в таблице. Для 3-й позиции Мифологического типа сумма рангов составит 3·4=12, а для 4-й: 4·19=76. Теперь определяем общую сумму рангов Мифологического типа: 2+12+12+76=102.

Таблица 9.19

Расчет ранговых сумм по четырем типам мужественности (n=31) для подсчета критерия χ^2_r.

Значение ранга	Типы мужественности							
	Мифологический		Национальный		Современный		Религиозный	
	f_{aj}	$f_{aj}\cdot r_j$	f_{aj}	$f_{aj}\cdot r_j$	f_{aj}	$f_{aj}\cdot r_j$	f_{aj}	$f_{aj}\cdot r_j$

1	2	2	19	19	7	7	3	3
2	6	12	4	8	10	20	11	22
3	4	12	7	21	12	36	8	24
4	19	76	1	4	2	8	9	36
Суммы рангов	102		52		71		85	

Сформулируем гипотезы.

H_0: Различия в позициях, которые занимают каждый из четырех типов мужественности, случайны,

H_1: Различия в позициях, которые занимают каждый из четырех типов мужественности, неслучайны. Определим эмпирическую величину χ^2_r по формуле:

$$\chi^2_r = \left[\frac{12}{n \cdot c \cdot (c+1)} \cdot \sum T_j^2\right] - 3 \cdot n \ (c+1)$$

где с - количество условий, в данном случае типов мужественности; *n* - количество испытуемых; T_j - суммы рангов по каждому из условий.

$$\chi^2_r = \left[\frac{12}{31 \cdot 4 \cdot (4+1)} \cdot \left(102^2 + 52^2 + 71^2 + 85^2\right)\right] - 3 \cdot 31 \cdot (4+1) = 26{,}11$$

Критические значения определяем по Табл. IX Приложения 1, поскольку при больших *n* χ^2_r имеет распределение, сходное с распределением χ^2, а существующие таблицы χ^2_r предназначены только для $n \leq 9$.

Количество степеней свободы определим так же, как мы это делали при расчете критерия χ^2:

$v = (k - 1)(c - 1) = (4 - 1)(4 - 1) = 3 \cdot 3 = 9$

При $v=9$ критические значения χ^2_r составляют:

$$\chi^2_{кр} = \begin{cases} 16{,}919 \ (p \leq 0{,}05) \\ 21{,}666 \ (p \leq 0{,}01) \end{cases}$$

$\chi^2_{эмп} = 26{,}11$

$\chi^2_{эмп} > \chi^2_{кр} \ (p \leq 0{,}01)$

Ответ: H_0 отвергается. Принимается H_1. Различия в позициях, которые занимает каждый из четырех типов мужественности, неслучайны ($p<0{,}01$). При этом на первом месте оказывается Национальный тип, на втором - Современный, на третьем - Религиозный и на четвертом - Мифологический тип. На Рис. 9.3. групповая система предпочтений представлена графически.

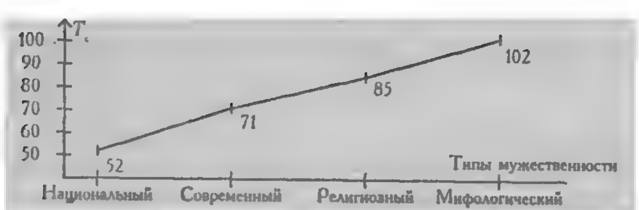


Рис. 9.3. Графики изменения ранговых сумм в последовательности: Национальный тип, Современный тип. Религиозный тип. Мифологический тип; меньшая сумма рангов указывает на большую предпочтительность типа, большая сумма - на меньшую предпочтительность

Итак, различия в ранговых местах каждого из рассматриваемых типов мужественности неслучайны. Наблюдается определенная групповая тенденция предпочтений. Судя по достаточно монотонному повышению кривой на Рис. 9.3, мы вряд ли можем говорить о резком преобладании какого-либо одного из двух типов мужественности. Для статистически достоверного ответа на этот вопрос необходимо сопоставить попарно все типы мужественности по схеме, использованной при решении Задачи 7.

9.5. Решения задач Главы 5

Решение задачи 9

Поскольку представлены данные по двум выборкам, мы выбираем критерий Фишера для оценки различий в процентных долях. Будем считать "эффектом" преобладание левого глаза. В исследовании Т.А. Доброхотовой и Н.Н. Брагиной высказывалось предположение о феномене предвосхищения у левшей, их способности к "зеркальному" отражению не только пространства, но и времени, выражающейся в прогностических возможностях и особого рода проницательности (Доброхотова Т.А., Брагина Н.Н., 1994). Интересно поэтому сопоставить выборки прямо по эффекту левшества.

Построим четырехпольную таблицу.

Таблица 9.20

Четырехпольная таблица для расчета критерия φ* при сопоставлении студентов-психологов (nj=14) и студентов-медиков (п2=100) по прицельной способности глаз

Группы	"Есть эффект" преобладание левого глаза		"Нет эффекта" преобладание правого глаза		Суммы
	Количество испытуемых	% доля	Количество испытуемых	% ДОЛЯ	
1 группа - студенты - психологи	6	42,9%	8	57,1%	14
2 группа - студенты - медики	19	19%	81	81%	100
Суммы	25		89		114

Сформулируем гипотезы.

H_0: Доля лиц с преобладанием левого глаза в группе студентов-психологов не больше, чем в группе студентов-медиков.

H_1: Доля лиц с преобладанием левого глаза в группе студентов-психологов больше, чем в группе студентов-медиков. По Табл. XII Приложения 1 определяем φ_1 и φ_2: $Ф_{1(42.9\%)}$=1.430; $ф_{2(19\%)}$=0,902 Подсчитываем эмпирическое значение φ *:

$$\varphi^*_{эмп} = (\varphi_1 - \varphi_2)\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (1{,}428 - 0{,}902) \cdot \sqrt{\frac{14 \cdot 100}{14 + 100}} = 1{,}84$$

Критические значения φ * нам известны:

$$\varphi^*_{кр} = \begin{cases} 1{,}64 \ (p \le 0{,}05) \\ 2{,}31 \ (p \le 0{,}01) \end{cases}$$

$\varphi^*_{эмп} > \varphi^*_{кр}$ ($p \le 0{,}05$)

Можно и более точно определить уровень значимости для $\varphi_{эмп}$=1,84: *p*=0,033.

Ответ: H_0 отклоняется. Доля лиц с преобладанием левого глаза в.группе студентов-психологов больше, чем в группе студентов-медиков (*p*=0,033).

Возможно, в данном исследовании произошло то, что называется "самоисполняющимися предсказаниями". Студенты-психологи перед началом опыта узнали об идее Брагиной и Доброхотовой о возможных прогностических способностях лиц с преобладанием левшества. Желание выявить у себя столь важные способности могло исказить результаты, несмотря на достаточную объективность карандашного теста.

Решение задачи 10

Поскольку сопоставляются две группы, мы применяем критерий φ * Фишера.

Вопрос 1: Можно ли считать, что милиционеры патрульно-постовой службы в большей степени склонны продолжить разговор с агрессором, чем другие граждане?

В данном случае нам легко определить, что будет критерием для разделения испытуемых на тех, у кого "есть эффект", и тех, у кого "нет эффекта".

Признак принимает всего два значения: разговор продолжен -разговор не продолжен. Ориентируясь на вопрос задачи, будем считать, что "эффект есть", если

разговор продолжен, и что "эффекта нет", если разговор не продолжен.

Создадим четырехклеточную таблицу.

Таблица **9.21**

Четырехклеточная таблица для сопоставления милиционеров (n_1=25) и гражданских лиц (n_2=25) по показателю продолжения разговора с агрессором

Группы	"Есть эффект" разговор продолжен		"Нет эффекта" разговор не продолжен		Суммы
	Количество испытуемых	% ДОЛЯ	Количество испытуемых	% ДОЛЯ	
1 группа - милиционеры	15	60%	10	40%	25
2 группа - гражданские лица	7	28%	18	72%	25
Суммы	22		28		50

Сформулируем гипотезы.

H_0: Доля лиц, продолживших разговор с агрессором, в группе милиционеров не больше, чем в группе гражданских лиц.

H_1: Доля лиц, продолживших разговор с агрессором, в группе милиционеров больше, чем в группе гражданских лиц.

Определим значения угла φ для каждой из сопоставляемых процентных долей:

$\varphi_{1(60\%)}$=1,772

$\varphi_{2(28\%)}$=1,115

Рассчитаем эмпирическое значение φ *:

$$\varphi^*_{кр} = (\varphi_1 - \varphi_2)\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (1{,}772 - 1{,}115)\sqrt{\frac{25 \cdot 25}{25 + 25}} = 2{,}32$$

Как мы помним, критические значения φ * для всех *щ,* п? равны:

$$\varphi^*_{кр} = \begin{cases} 1{,}64\ (p \le 0{,}05) \\ 2{,}31\ (p \le 0{,}01) \end{cases}$$

$\varphi^*_{эмп} > \varphi^*_{кр}$ $(p<0{,}01)$

Ответ: H_0 отвергается. Принимается H_1. Доля лиц, продолживших разговор с агрессором, в группе милиционеров больше, чем в группе гражданских лиц (p<0,01).

Вопрос 2: Можно ли утверждать, что милиционеры склонны отвечать агрессору более примирительно, чем гражданские лица?

Ориентируясь на вопрос задачи, будем считать, что "эффект есть", если испытуемый дал неагрессивный, примирительный ответ, и что "эффекта нет", если испытуемый дал агрессивный ответ.

Создавая четырехклеточную таблицу, мы помним о том, что объемы выборок несколько сократились, поскольку в них были испытуемые, уклонившиеся от продолжения разговора с агрессором.

Таблица 9.22

Четырехклеточная таблица для сопоставления долей неагрессивных ответов в группах милиционеров и гражданских лиц

Группы	"Есть эффект" неагрессивный ответ		"Нет аффекта" агрессивный ответ		Суммы
1 группа - милиционеры	10	(66,7%)	5	(33,3%)	15

2 группа - гражданские лица	3 (42,9%)	4 (57,1%)	7
Суммы	13	9	22

Мы можем далее использовать угловое преобразование Фишера, то есть критерий φ *, поскольку ни одна из сопоставляемых долей не равна 0. При уменьшении выборок угроза появления нуля в какой-либо из ячеек, естественно, возрастает, поэтому контроль необходим (см. ограничение 1).

Сформулируем гипотезы. H_0: Доля неагрессивных ответов в группе милиционеров не больше, чем

в группе гражданских лиц. H_1: Доля неагрессивных ответов в группе милиционеров больше, чем в

группе гражданских лиц. Определим значения φ_1 и φ_2 и рассчитаем эмпирическое значение φ *:

$\varphi_{1(66,7\%)}=1,911$

$\varphi_{2(42,9\%)}=1,428$

$$\varphi^*_{кр} = (\varphi_1 - \varphi_2)\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (1,911 - 1,428)\sqrt{\frac{15 \cdot 3}{15+3}} = 1,06$$

$$\varphi^*_{кр} = \begin{cases} 1,64\ (p \le 0,05) \\ 2,31\ (p \le 0,01) \end{cases}$$

$\varphi^*_{вмп} < \varphi^*_{кр}$

Ответ: H_0 принимается. Доля неагрессивных ответов в группе милиционеров не больше, чем в группе гражданских лиц. Как свидетельствует Табл.XIII Приложения 1, полученное значение меньше даже критического значения, соответствующего уровню значимости $p<0,10$.

Вспомним, однако, что мы включили в четырехклеточную таблицу только тех испытуемых обеих групп, которые вступили в разговор с агрессором. Если же подсчитывать доли агрессивных и неагрессивных реакций по отношению к общему количеству испытуемых в данной выборке, то результат получится достоверным ($\varphi^*_{эмп}=2,34$; $p<0,01$).

Именно так и поступил автор данного исследования. Это позволило ему сделать вывод, что милиционеры предпочитают неагрессивные ответы, поскольку, по-видимому, это дает время и возможность оценить степень опасности нападающего субъекта для окружающих. Срабатывает профессиональный инстинкт (Кузнецов А.А., 1991).

Решение задачи 11

Поскольку сравниваются две выборки, выбираем критерий φ * Фишера. Однако "на глаз" трудно решить, какая команда врачей должна считаться большей по составу, а какая - меньшей. Нам необходимо найти точку, в которой накапливаются максимальные различия между двумя распределениями, для того, чтобы применение критерия φ * было максимально эффективным. Для этого вначале используем алгоритм определения максимальной разности между накопленными частостями, используемый в критерии λ Колмогорова-Смирнова (Алгоритм 15).

Результаты применения алгоритма представлены в Табл. 9.23

Таблица 9.23

Выявление точки максимальной разности между эмпирическими распределениями "количества партнеров у врачей с фондами (n_1=49) и врачей без фондов (n_2=28)

количество партнеров		Эмпирические частоты		Эмпирические частости		Накопленные эмпирические частости		Разность
		f_1	f_2	f^*_1	f^*_2	$\sum f^*_1$	$\sum f^*_2$	$(\sum f^*_1 - \sum f^*_2)$
1	2 и менее партнеров	2	15	0,041	0,536	0,041	0,536	0,495
2	3-4 партнера	6	5	0,122	0,179	0,163	0,715	0552
3	5-6 партнеров	27	8	0,551	0,286	0,714	1,000	0,286
4	7 и более партнеров	14	0	0,286	0	1,000	1,000	0
Суммы		49	28	1,000	1,000			

Как видно из Табл. 9.23, максимальная разность накопленных частостей падает на 2-й разряд (3-4 партнера). Поскольку вопрос в задаче касается предположения о том, что в приемных с фондами работают большие по составу команды врачей, чем в приемных без фондов, будем считать, что если партнеров более 4-х, то "эффект есть", а если партнеров 4 и менее, то "эффекта нет". Построим соответствующую четырехклеточную таблицу и определим % доли "эффекта" в каждой из двух выборок.

Таблица 9.24

Четырехклеточная таблица для подсчета критерия φ * при сопоставлении выборок врачей с фондами (n_1=49) и врачей без фондов *(n_2=28)* по признаку количества партнеров

Группы	"Есть эффект" более 4 партнеров		"Нет эффекта" не более 4 партнеров		Суммы
1 группа - врачи с фондами	41	(83,7%)	8	(16,3%)	49
2 группа - врачи без фондов	8	(28,6%)	20	(71,4%)	28
Суммы	49		28		77

Сформулируем гипотезы.

H_0: Доля лиц, имеющих более 4-х партнеров, в выборке врачей с фондами не больше, чем в выборке врачей без фондов.

H_1: Доля лиц, имеющих более 4-х партнеров, в выборке врачей с фондами больше, чем в выборке врачей без фондов. По Табл. XII Приложения 1 определяем углы φ:

$\varphi_{1(83,7\%)}$=2,310

$\varphi_{2(28,6\%)}$=1,129

Рассчитаем эмпирическое значение критерия φ *:

$$\varphi^*_{кр} = (\varphi_1 - \varphi_2)\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (2{,}310 - 1{,}129)\sqrt{\frac{49 \cdot 28}{49 + 28}} = 4{,}99$$

$$\varphi^*_{кр} = \begin{cases} 1{,}64 \ (\rho \le 0{,}05) \\ 2{,}31 \ (\rho \le 0{,}01) \end{cases}$$

По Табл. XIII Приложения 1 определяем, какому уровню значимости соответствует эта величина φ *: p<0,001.

$\varphi^*_{эмп} > \varphi^*_{кр}$ (p<0,001)

Ответ: H_0 отклоняется. Принимается H_1. Доля лиц, имеющих более 4-х партнеров, в выборке врачей с фондами больше, чем в выборке врачей без фондов (p<0,001).

Фирма с высокой степенью уверенности может ориентироваться на эту тенденцию в построении своей стратегии продвижении товара. Но то, как она будет ее учитывать, уже выходит за рамки данной статистической задачи.

Решение задачи 12

Обследована одна выборка испытуемых, поэтому останавливаем выбор на биномиальном критерии m. В параграфе 4.2, посвященному методу χ^2, эта задача предположительно должна была решаться с помощью критерия χ^2. Однако, поскольку количество наблюдений $n<300$, а вероятность выбора каждой из дорожек при равновероятном выборе составляет ½, т.е. P=Q=0,50, мы можем воспользоваться биномиальным критерием, который несравненно проще в использовании, чем критерий χ^2. Воспроизведем таблицу частот.

Таблица 9.25

Эмпирические частоты выбора правой и левой симметричных дорожек (n=70)

Выбрана правая дорожка	Выбрана левая дорожка	Суммы
51	19	70

Сформулируем гипотезы.

H_0: Частота выбора правой дорожки не превышает частоты, которая соответствует вероятности случайного выбора.

H_1: Частота выбора правой дорожки превышает частоту, которая соответствует вероятности случайного выбора.

Определим теоретическую частоту выбора одной из дорожек при случайном выборе:

$f_{теор}=n\cdot p=70\cdot 0{,}50=35$

Поскольку $f_{эмп} > f_{теор}$, используем биномиальный критерий m, а не его "зеркальное отражение" (критерий знаков G).

По Табл. XIV Приложения 1 определяем критические значения m для n=70:

$$m_{кр}=\begin{cases}43\ (\rho\le 0{,}05)\\46\ (\rho\le 0{,}01)\end{cases}$$

$m_{эмп}=f_{эмп}=51$

$m_{эмп}>m_{кр}$.

Ответ: H_0 отклоняется. Принимается H_1. Частота предпочтения правой дорожки превышает частоту, которая соответствует вероятности случайного выбора (p<0,01).

Наблюдатель может обоснованно утверждать, что из данных двух симметричных дорожек чаще выбирается правая. Чем это объясняется— уже другой вопрос, выходящий за рамки задачи (см. п. 4.2).

БИБЛИОГРАФИЯ

ЗАКЛЮЧЕНИЕ

Психология - молодая наука. Все, что молодо, сохраняет гибкость и способность к изменению.

Наша наука допускает поэтому существование самых разных моделей высокого профессионализма. Среди моих знакомых есть, например, коллега, который фактически никогда не узнавал испытуемых, с которыми мы сталкивались во дворе того предприятия, где совместно проводили многодневное исследование, и даже не здоровался с ними. Вместе с тем, он был способен по протоколу эксперимента или по профилю MMPI дать столь глубокий и точный портрет совершенно неизвестного ему лица, что это поражало и вызывало неизменное удивление. Знание среднестатистических закономерностей, по-видимому, может компенсировать отсутствие личного впечатления от человека.

С другой стороны, я сталкиваюсь и с другими коллегами, которые полагаются скорее на личное восприятие человека и не считают среднестатистические закономерности определяющими в исследовании отдельного человека. Их профессиональные возможности тоже бывают поразительными.

Психологу никогда не бывает скучно, потому что он всегда изучает и исследует - людей, ситуации, самого себя. Он постоянно ищет свой путь в выявлении новых закономерностей и фактов.

Методы математической статистики могут оказать на этом пути неоценимую помощь, но они - лишь средство, которое не должно заслонять собою цель. Необходимо помнить, что достоверная статистическая тенденция - это все же не психологическая закономерность, а выпадающие из общей картины индивидуальные значения - не артефакт, а отражение, быть может, закономерности более высокого порядка, чем те, что выявляются с помощью математических методов.

Если продолжить аналогию С.Стивенса с веревочной лестницей, то мы используем веревочную лестницу, чтобы подняться наверх, хотя знаем, что и без нее можем летать. Главное — чтобы из-за привязанности к веревочной лестнице мы не утратили этой способности к полету.

1 *Алимов Ю.И.* Альтернатива методу математической статистики. М.: Знание, 1986. 64 с.

2 *Ананьев Б.Г.* Человек как предмет познания. Л.:ЛГУ, 1969. 339 с.

3 *Ананьев Б.Г.* О методах современной психологии // Психодиагностические методы (в комплексном лонгитюдном исследовании студентов). Л.: ЛГУ, 1976. С. 13-35.

4 *Андреенков ВТ., Аргунова К.Д. и др.* Математические методы анализа и интерпретация социологических данных. // Под ред. В.Г. Андреенкова, Ю.Н. Толстовой. М.: Наука, 1989. 171 с.

5 *Артемьева Е.Ю., Мартынов Е.М.* Вероятностные методы в психологии. М: МГУ, 1985. 206 с.

6 *Ашмарин И.П., Васильев Н.Н., Амбросов ВЛ.* Быстрые методы статистической обработки и планирование экспериментов. Л.: ЛГУ, 1974. 76 с.

7 *Бадасова Г.А.* Личностные факторы суггестора, способствующие внушающему воздействию. Дипломная работа выпускницы специального факультета социальной психологии СПбГУ. СПб, 1994. 75 с.

8 *Бергер Н.А., Логинова Н.А.* К проблеме соотношения некоторых содержательных и структурных характеристик интеллекта (по методике Векслера).// Современные психолого-педагогические проблемы высшей школы. Л.: ЛГУ,

1974. С. 63-66.

9 *Берн Э.* Игры, в которые играют люди. Психология человеческих взаимоотношений; Люди, которые играют в игры. Психология человеческой судьбы. / Пер. с англ. // Общ. ред. М.С. Мацковского. СПб.: Лениздат, 1992. 400 с.

10 *Большее Л.Н., Смирнов Н.В.* Таблицы математической статистики. М.: Наука. Главн. редакция физико-математ. литературы, 1983. 416 с.

И *Бурлачук Л.Ф., Морозов СМ.* Словарь-справочник по математической диагностике. Киев.: Наук, думка, 1989. 200 с.

12 *Ван дер Варден Б.Л.* Математическая статистика. М., 1960. 434 с.

13 *Гайда В.К., Захаров В.П.* Психологическое тестирование. Учебное пособие. Л.: ЛГУ, 1982.101 с.

14 *Ганзен ВЛ,. Балин В.Д.* Теория и методология психологического исследования. Практическое руководство. СПб.: СПбГУ, 1991. 74 с.

15 *Геодакян В А.* Дифференциальная смертность и норма реакции мужского и женского пола. Онтогенетическая и филогенетическая пластичность. // Журнал общей биологии, 1974, т.35, №3. С. 376-385.

16 *Геодакян ВЛ.* Асинхронная асимметрия (половая и латеральная дифференциация — следствие асинхронной эволюции). // Журнал ВНД, 1993, т.43. Вып.З. С. 543-561.

17 *Гласе Дж., Стенли Дж.* Статистические методы в педагогике и психологии. / Пер. с англ. под общ. ред. Ю.П. Адлера. М.: Прогресс,- 1976. 495 с.

18 *Гоголь Н.В,* Избранные произведения. М.: ДетГИЗ, 1959, С. 473-500.

19 *Грекова И.* Методологические особенности прикладной математики на современном этапе ее развития. // Вопросы философии, 1976, №6, С. 104-114.

20 *Гублер Е.В.* Вычислительные методы анализа и распознавания патологических последствий. Л.: Медицина, 1978. 296 с.

21 *Гублер Е.В., Генкин А.А.* Применение непараметрических критериев статистики в медико-биологических исследованиях. Л.: Медицина, 1973. 142 с.

22 *Девятка И.Ф.* Диагностическая процедура в социологии. Очерки истории и теории. М.: Наука, 1993. 173 с.

23 *Дворяшина М.Д., Пехлецкий И.Д.* Основные математические процедуры психодиагностического исследования.// Психодиагностические методы (в комплексном лонгитюдном исследовании студентов). Л.: ЛГУ, 1976. С. 35-51.

24 *Доброхотова ТА., Брашна Н.Н.,* Левши. М.: Книга, 1994. 230 с.

25 *Езекизл М., Фокс К.А.* Методы анализа корреляций и регрессий (линейных и криволинейных).// Пер. с англ. Л.С. Кучаева. М.: Статистика, 1966. 559 с.

26 *Захаров В.П.* Применение математических методов в социально-психологических исследованиях. Учебное пособие. Л.: ЛГУ, 1985. 64 с.

27 *Ивантер Э.В., Коросов А.В.* Основы биометрии: Введение в статистический анализ биологических явлений и процессов. Учебное пособие. Петрозаводск: ПГУ, 1992. 163 с.

28 *Ильин Е.П.* Психофизиология физического воспитания. Деятельность и состояния. Учебное пособие для студентов факультетов физического воспитания педагогических институтов. М.: Просвещение, 1980. 199 с.

29 *Ильина М.Н.* Способность к проявлению терпения при мышечном утомлении как отражение общего волевого фактора. / Психомоторика. Сборник ученых трудов. // Под ред. Б.А. Ашмарина и проф Е.П. Ильина (научн. ред.). Л.: ЛГПИ, 1976. С. 49-50.

30 *Кендалл М.Дж., Стюарт А.* Статистические алгоритмы в социологических исследованиях. Новосибирск: Наука, 1985. 207 с.

31 *Кенуй М.Г.* Быстрые статистические вычисления. Упрощенные методы оценивания и проверки. / Пер. с англ. и предисловие Д.А. Астринского. М.: Статистика, 1979. 69 с.

32 *Королькова Н.А.* Возможности психологической коррекции у болезненных детей. Дипломная работа выпускницы кафедры социальной психологии факультета психологии СПбГУ. СПб., 1994. 72 с.

33 *Кузнецов С~А.* Стили реагирования на вербальную агрессию. Дипломная работа выпускника кафедры социальной психологии факультета психологии СПбГУ. СПб., 1991. 33 с.

34 *Кулева Е.Б.* Влияние традиционных и православных текстов внушения на процесс аутогенной тренировки. Дипломная работа выпускницы кафедры социальной психологии факультета психологии СПбГУ. СПб., 1990. 45 с.

35 *Курочкин М.А., Сидоренко ЕЈЗ., Чураков ЮА. (Kurochkin M., Churakov U., Sidorenko E.).* Opportunities for Leadership in Healthcare. General Practicin-ers Research Project for Lilly Industries. Manchester. Manchester Business School, 1992. 22 p.

36 *Лашков К.В., Поляков Л.Е.* Непараметрические методы медико-статистических исследований. / Методологические вопросы санитарной статистики. Ученые записки по статистике, т.IX. М.: Наука, 1965. С. 136-184.

37 *Логвиненко А.Д.* Измерения в психологии М.: МГУ.1993. 480 с.

38 Математические методы анализа и интерпретация социологических данных. // Отв. ред. В.Г. Андреенков, Ю.Н. Толстова. М.: Наука, 1989. 171 с.

39 Математические методы психолого-педагогических исследований. Методические рекомендации. СПб.: Образование, 1994. 28 с.

40 *Мельников В.М., Ямпольский Л.Т.* Введение в экспериментальную психологию личности. Учебное пособие для слушателей ИПК преподавателей педагогических дисциплин университетов и педагогических институтов. М.: Просвещение, 1985. 319 с.

41 Методы современной биометрии. М.: МГУ, 1978. С.108-179.

42 *Митропольский А.К.* Техника статистических вычислений. М.: Наука, Главная редакция физико-математической литературы., 1971. 576 с.

43 *Михеев В.Н.* Методика получения и обработки экспериментальных данных в психолого-педагогических исследованиях. М.: УДН, 1986. 84 с.

44 *Налимов В.В.* Теория эксперимента. М.: Наука, 1975. 207 с.

45 *Налимов В.В., Голикова Т.Н.* Логические основания планирования эксперимента. Изд. 2-е. М.:-Металлургия, 1981. 152 с.

46 *Нискина Н.П.* Непараметрические методы математической статистики и решение задач проверки гипотез./ Проблемы компьютеризации и статистики в прикладных науках. Сборник трудов. М.: ВНИИСИ, 1990. С. 73-89.

47 *Носенко И.А.* Начала статистики для лингвистов. М.: Высшая школа, 1981. 157 с.
48 Оуэн *Д.Б.* Сборник статистических таблиц. / Пер. с англ. Л.Н. Большева и В.Ф. Котельниковой. Изд. 2-е, исправл. М.: Вычислительный центр АН СССР. 1973. 586 с.
49 *Паповян С.С.* Математические методы в социальной психологии. М.: Наука, 1983. 343 с.
50 *Плохинский Н~А.* Дисперсионный анализ. / Под ред. чл.-корр. АН СССР Н.П. Дубинина. Новосибирск: Сиб. Отд. АН СССР, 1960. 124 с.
51 *Плохинский Н.А.* Биометрия. 2-е изд. М.: МГУ, 1970. 368 с.
52 *Пуни А.Ц.* Психологические основы волевой подготовки в спорте. Учебное пособие. Л.: ГИФК, 1977. 48 с.
53 *Пустыльник Е.И.* Статистические методы анализа и обработки наблюдений. М,: Наука, 1968. 185 с.
54 *Рахова М.Э.* Личностная предрасположенность к определенным видам страха. Дипломная работа выпускницы кафедры социальной психологии факультета психологии СПбГУ. СПб., 1994. 54 с.
55 *Роджерс К.* Взгляд на психотерапию. Становление человека. / Пер. с англ. //Общ. ред. и предисл. Е.И.Исениной. М.: Прогресс, Универс, 1994. 480 с.
56 *Рунион Р.* Справочник по непараметрической статистике. М.: Финансы и статистика, 1982. 198 с.
57 *Сидоренко (Маркова) Е.В.* Связь мотивации достижения с индивидными и личностными свойствами / Вопросы экспериментальной и прикладной психологии. Сборник аспирантских работ. Л.: ЛГУ, 1980. Деп. в ВНТИ №435-80 от 7 февр. 1980. С. 64-72
58 *Сидоренко (Маркова) Е.В.* Исследование психодиагностических возможностей проективной методики Хекхаузена. / Личность в системе коллективных отношений. Тезисы докладов Всесоюзной конференции в г.Курске. Курск: 1980. С. 43-45
59 *Сидоренко (Маркова) Е.В.* Мотивационно>волевые особенности личности как фактор успешной деятельности. Дисс. на соискание учен. степ. канд. психол. наук. Л.: ЛГУ, 1984. 262 с.
60 *Сидоренко (Маркова) Е.В.* Психодраматический и недирективный подходы в групповой работе с людьми. Методические описания и комментарии. СПб.: Центр психологической поддержки учителя, 1992. 72 с.
61 *Сидоренко Е.В.* Экспериментальная групповая психология. Комплекс "неполноценности" и анализ ранних воспоминаний в концепции Альфреда Адлера. Учебное пособие. СПб.: СПбГУ, 1993. 152 с.
62 *Сидоренко Е.В.* Опыты реоритационного тренинга. СПб.: Институт тренинга, 1995. 248 с.
63 *Сидоренко Е.В., Дсрманова И.Б., Анисимова О.М., Витенберг Е.В.. Шульга А.П.* Разработка методики отбора и подготовки кадров в представительные органы муниципальной власти. СПб.: Гуманистический и политологический Центр "Стратегия", 1994. 26 с.
64 *Сочивко Д.В., Якунин В.А.* Математические модели в психолого-педагогических исследованиях. Учебное пособие. Л.: ЛГУ, 1988. 68 с.
65 Справочник по прикладной статистике. В 2-х т. Т.2 / Пер. с англ. под ред. Э.Ллойда, У. Ледермана, С.А. Айвазяна, Ю.Н. Тюрина. М.: Финансы и статистика, 1990. 526 с.
66 *Стан Н.В.* Социально-психологическое исследование стереотипов мужественности. Дипломная работа выпускницы кафедры социальной психологии факультета психологии СПбГУ. СПб., 1992. 58 с.
67 *Стивене С.* Математика, измерение и психофизика // Экспериментальная психология (Под ред. С.С. Стивенса). // Пер. с англ под ред. действ, чл. АМН СССР П.К. Анохина, докт. пед. наук В.А. Артемова. М.: Иностранная литература, 1960. т.1. С. 19-92.
68 *Суходольский Г.В.* Основы математической статистики для психологов. Л.: ЛГУ, 1972. 428 с.
69 *Суходольский Г.В.* Математике-психологические модели деятельности. СПб.: Петрополис, 1994. 64 с.
70 *Тлегенова Г.А.* Влияние агрессивности на проксемические характеристики невербального поведения. Дипломная работа выпускницы кафедры социальной психологии факультета психологии СПбГУ. СПб., 1990. 28 с.
71 *Толстова Ю.Н.* Логика математического анализа социологических данных. М.: Наука, 1991.112 с.
72 *Тюрин Ю.Н.* Непараметрические методы статистики. М.: Знание, 1978. 64 с.
73 *Тюрин Ю.Н., Макаров А.А,* Анализ данных на компьютере. // Под ред. В.В. Фигурнова.

М.: Финансы и статистика, 1995. 384 с.
74 *Урбах В.Ю.* Математическая статистика для биологов и медиков. М.: Академия наук СССР, 1963. 323 с.
75 *Урбах В.Ю.* Биометрические методы. Статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине. М.: Наука, 1964. 415 с.
76 *Урбах В.Ю.* Статистический анализ в биологических и медицинских исследованиях. М.: Медицина, 1975. 295 с.
77 *Фелингер А.Ф.* Статистические алгоритмы в социологических исследованиях. Новосибирск: Наука, 1985. 385 с.
78 *Холлендер М. Вулф Д.А.* Непараметрические методы статистики. / Пер. с англ. под ред. Ю.П. Адлера и Ю.Н. Тюрина М.: Финансы и статистика, 1983. 518 с.
79 *Чиркина Р.Т.* Психодинамические факторы памяти. Дипломная работа выпускницы кафедры социальной психологии факультета психологии СПбГУ. СПб., 1995. 80 с.
80 *Шеффе Г.* Дисперсионный анализ. М.: Наука, 1980. 512 с.
81 *Caltell R.B.. Eber H.W..Tatsuoka M.M.* Handbook (or the Sixteen Personality Factor Questionnaire (16PF) Champaign, Illinois: Institute for Personality and Ability Testing, 1970.
82 *Fisher R.A.* Statistical Methods for Research Workers. N.Y., 1925. 356 p.
83 *Fridman M.* The Use of Ranks to Avoid the Assumption of Normality in the Analysis of Variance. Journal of American Statistical Association, 1937. N 32. P. 675-701.
84 *Greene /., D'Olivera M.* Learning to Use Statistical Tests in Psychology: a Student's Guide. Milton Keynes Philadelphia, Open University Press, 1989. 180 p.
85 *Hall E.T..* The Silent Language. Greenwich (Conn.): Fawcett, 1970. 192 p.
86 Warns *K.A.. Morrow K..B.* Differential Effects of Birth Order and Gender on Perceptions of Responsibility and Dominance.// Individual Psychology: The Jour-nal of Adlerian Theory, Research and Practice, 1992. Vol.48. N.I. P.109-118.
87 *Klar H.* Lusher-test Copyriht und Bestellugen bei Test. Basel (Schweiz.): Verlag, Graphshe Gestaltung und Brack: Rentch AG, Zrimbach-Olten, 1974. P. 103-126
88 *Krauth }.* Distribution - Free Statistics: An Application-Oriented Approach. Amsterdam, 1988. 381 p.
"89 *Kruskal W.H., Wallis W.A.* Use of Ranks iii One-Criterion Variance Analysis. Journal of the American Statistical Association, 47 (1952), P. 583-621, and 48
(1953), P. 907-911.
90 *Kuril AX., Mayo S.T.* Statistical Methods in Education and Psychology. N.Y., etc.: Springer-Verlag, 1979. 538 p.
91 Louie *A.D.* New Developements in Statistics for Psychology and the Social Sciences. London and N.Y., The British Psychological Society and Medcene, 1986. 17Я *n*
92 *Manaslcr* С.У., Corsini *R.Y.* Individual Psychology. Itasca: F.E. Peacock Publishers, Inc., 1982. 322 p.
93 Mathematical Psychology. / Ed. by R. Duncan Luce, et al. N.Y.-London, 1963.
95 *McCati R.* Fundamental Statistics for Psychology. N.Y., Harcont, Brace E.
World, 1970. 418 p.
96 *McClelland D.C., Winter D.G.* Motivating Economic Achievement. N.Y. - London: The Free Press, 1971. 415 p.
97 *McNemar* Q. Psychological Statistics. N.Y. - London: Wiley and sons, 196Э.
98 Moreno *J.L.* Who Shall Survive? A New Approach to the Problem of Human Interrelations. Wasington: Nervous and Mental Disease Publishing Co., 1934.
99 *Moreno J.L.* Sociometry, Experimental Method and the Science of Society. Beacon (N.Y.): Beacon House, 1951.
100 *Rogers* C. O. Becomiirg a Person. A Therapist's View of Psychotherapy. Boston: Houghton Mifflin Co, 1961. 420 p.
101 Rovine M.J., von Eye A. Applied Computational Statistics in Longitudinal Research. Boston, etc.: Acad.Press. 1991. 237 p.
102 *Welkowitz /., Ewen R.B., Cohen J.* Introductory Statistics for the Behavioral bciences. - 3-d Ed. - N.Y. etc., Academic Press, 1982. 372 p
103 Шсохоп F Inchvidual Comparisons by Ranking Methods. - Biometrics Bulletin, iy^D. JN21, I-'. oU-o3

104 *Wolpe J. (with D. Wolpe)* Our Useless Fears. Boston: Houghton Mifflin, 1981. lol p.

ПРИЛОЖЕНИЕ 1
ТАБЛИЦЫ КРИТИЧЕСКИХ ЗНАЧЕНИЙ

Таблица 1

Критические значения критерия Q Розенбаума для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ (по Гублеру Е.В., Генкину А.А., 1973)

Различия между двумя выборками можно считать достоверными ($\rho \leq 0,05$), если $Q_{эмп}$ равен или выше критического значения $Q_{0,05}$, и тем более достоверными ($\rho \leq 0,01$), если $Q_{эмп}$ равен или выше критического значения $Q_{0,01}$.

n	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
								ρ=0,05								
11	6															
12	6	6														
13	6	6	6													
14	7	7	6	6												
15	7	7	6	6	6											
16	8	7	7	7	6	6										
17	7	7	7	7	7	7	7									
18	7	7	7	7	7	7	7	7								
19	7	7	7	7	7	7	7	7	7							
20	7	7	7	7	7	7	7	7	7	7						
21	8	7	7	7	7	7	7	7	7	7	7					
22	8	7	7	7	7	7	7	7	7	7	7	7				
23	8	8	7	7	7	7	7	7	7	7	7	7	7			
24	8	8	8	8	8	8	8	8	8	8	7	7	7	7		
25	8	8	8	8	8	8	8	8	8	7	7	7	7	7	7	
26	8	8	8	8	8	8	8	8	8	8	7	7	7	7	7	7
								ρ=0,01								
11	9															
12	9	9														
13	9	9	9													
14	9	9	9	9												
15	9	9	9	9	9											
16	9	9	9	9	9	9										
17	10	9	9	9	9	9	9									
18	10	10	9	9	9	9	9	9								
19	10	10	10	9	9	9	9	9	9							
20	10	10	10	10	9	9	9	9	9	9						
21	11	10	10	10	9	9	9	9	9	9	9					
22	11	11	10	10	10	9	9	9	9	9	9	9				
23	11	11	10	10	10	10	9	9	9	9	9	9	9			
24	12	11	11	10	10	10	10	9	9	9	9	9	9	9		
25	12	11	11	10	10	10	10	10	9	9	9	9	9	9	9	
26	12	12	11	11	10	10	10	10	10	9	9	9	9	9	9	9

Таблица II

Критические значения критерия U Манна-Уитни для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ (по Гублеру Е.В., Генкину А.А., 1973)

Различия между двумя выборками можно считать значимыми ($\rho < 0,05$), если $U_{эмп}$ ниже или равен $U_{0,05}$, и тем более достоверными ($\rho < 0,01$), если $U_{эмп}$ ниже или равен $U_{0,01}$.

n_1 / n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	$\rho = 0,05$																		
3	-	0																	
4	-	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	6	9	12	15	18	21											
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	$\rho = 0,01$																		
5	-	-	0	1															
6	-	-	1	2	3														
7	-	0	1	3	4	6													
8	-	0	2	4	6	7	9												
9	-	1	3	5	7	9	11	14											
10	-	1	3	6	8	11	13	16	19										
11	-	1	4	7	9	12	15	18	22	25									
12	-	2	5	8	11	14	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	26	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Таблица II. Продолжение

n_1 / n_2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	$\rho = 0,05$																	
21	19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154
22	20	28	36	44	52	60	69	77	85	94	102	111	119	128	136	145	154	162
23	21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170
24	22	31	39	48	57	66	75	85	94	103	113	122	131	141	150	160	169	179
25	23	32	41	50	60	69	79	89	98	108	118	128	137	147	157	167	177	187
26	24	33	43	53	62	72	82	93	103	113	123	133	143	154	164	174	185	195
27	25	35	45	55	65	75	86	96	107	118	128	139	150	160	171	182	193	203
28	26	36	47	57	68	79	89	100	111	122	133	144	156	167	178	189	200	212
29	27	38	48	59	70	82	93	104	116	127	139	150	162	173	185	196	208	220
30	28	39	50	62	73	85	96	108	120	132	144	156	168	180	192	204	216	228
31	29	41	52	64	76	88	100	112	124	137	149	161	174	186	199	211	224	236
32	30	42	54	66	78	91	103	116	129	141	154	167	180	193	206	219	232	245
33	31	43	56	68	81	94	107	120	133	146	159	173	186	199	213	226	239	253
34	32	45	58	71	84	97	110	124	137	151	164	178	192	206	219	233	247	261
35	33	46	59	73	86	100	114	128	142	156	170	184	198	212	226	241	255	269
36	35	48	61	75	89	103	117	132	146	160	175	189	204	219	233	248	263	278
37	36	49	63	77	92	106	121	135	150	165	180	195	210	225	240	255	271	286
38	37	51	65	79	94	109	124	139	155	170	185	201	216	232	247	263	278	294
39	38	52	67	82	97	112	128	143	159	175	190	206	222	238	254	270	286	302
40	39	53	69	84	100	115	131	147	163	179	196	212	228	245	261	278	294	311
	$\rho = 0,01$																	
21	10	16	22	29	35	42	49	56	63	70	77	84	91	98	105	113	120	127
22	10	17	23	30	37	45	52	59	66	74	81	89	96	104	111	119	127	134
23	11	18	25	32	39	47	55	62	70	78	86	94	102	109	117	125	133	141
24	12	19	26	34	42	49	57	66	74	82	90	98	107	115	123	132	140	149
25	12	20	27	35	44	52	60	69	77	86	95	103	112	121	130	138	147	156
26	13	21	29	37	46	54	63	72	81	90	99	108	117	126	136	145	154	163
27	14	22	30	39	48	57	66	75	85	94	103	113	122	132	142	151	161	171
28	14	23	32	41	50	59	69	78	88	98	108	118	128	138	148	158	168	178
29	15	24	33	42	52	62	72	82	92	102	112	123	133	143	154	164	175	185
30	15	25	34	44	54	64	75	85	95	106	117	127	138	149	160	171	182	192
31	16	26	36	46	56	67	77	88	99	110	121	132	143	155	166	177	188	200
32	17	27	37	47	58	69	80	91	103	114	126	137	149	160	172	184	195	207
33	17	28	38	49	60	72	83	95	106	118	130	142	154	166	178	190	202	214
34	18	29	40	51	62	74	86	98	110	122	134	147	159	172	184	197	209	222
35	19	30	41	53	64	77	89	101	114	126	139	152	164	177	190	203	216	229
36	19	31	42	54	67	79	92	104	117	130	143	156	170	183	196	210	223	236
37	20	32	44	56	69	81	95	108	121	134	148	161	175	189	202	216	230	244
38	21	33	45	58	71	84	97	111	125	138	152	166	180	194	208	223	237	251
39	21	34	46	59	73	86	100	114	128	142	157	171	185	200	214	229	244	258
40	22	35	48	61	75	89	103	117	132	146	161	176	191	206	221	236	251	266

Таблица II. Продолжение

n_1 / n_2	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	p=0,05																		
21																			
22	171																		
23	180	189																	
24	188	198	207																
25	197	207	217	227															
26	206	216	226	237	247														
27	214	225	236	247	258	268													
28	223	234	245	257	268	279	291												
29	232	243	255	267	278	290	302	314											
30	240	252	265	277	289	301	313	326	338										
31	249	261	274	287	299	312	325	337	350	363									
32	258	271	284	297	310	323	336	349	362	375	389								
33	266	280	293	307	320	334	347	361	374	388	402	415							
34	275	289	303	317	331	345	359	373	387	401	415	429	443						
35	284	298	312	327	341	356	370	385	399	413	428	442	457	471					
36	292	307	322	337	352	367	381	396	411	426	441	456	471	486	501				
37	301	316	332	347	362	378	393	408	424	439	454	470	485	501	516	531			
38	310	325	341	357	373	388	404	420	436	452	467	483	499	515	531	547	563		
39	318	335	351	367	383	399	416	432	448	464	481	497	513	530	546	562	579	595	
40	327	344	360	377	394	410	427	444	460	477	494	511	527	544	561	578	594	611	628
	p=0,01																		
21																			
22	142																		
23	150	158																	
24	154	166	174																
25	163	174	183	192															
26	173	182	191	201	210														
27	180	190	200	209	219	229													
28	188	198	208	218	229	239	249												
29	196	206	217	227	238	249	259	270											
30	203	214	225	236	247	258	270	281	292										
31	211	223	234	245	257	268	280	291	303	314									
32	219	231	242	254	266	278	290	302	314	326	338								
33	227	239	251	263	276	288	300	313	325	337	350	362							
34	234	247	260	272	285	298	311	323	336	349	362	375	387						
35	242	255	268	281	294	308	321	334	347	360	374	387	400	413					
36	250	263	277	290	304	318	331	345	358	372	386	399	413	427	440				
37	258	271	285	299	313	327	341	355	370	384	398	412	426	440	454	468			
38	265	280	294	308	323	337	352	366	381	395	410	424	439	453	468	482	497		
39	273	288	303	317	332	347	362	377	392	407	422	437	452	467	482	497	512	527	
40	281	296	311	326	342	357	372	388	403	418	434	449	465	480	495	511	526	542	557

Таблица II. Продолжение

n_1 / n_2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	p=0,05																	
41	40	55	70	86	102	118	135	151	168	184	201	218	234	251	268	285	302	319
42	41	56	72	88	105	121	138	155	172	189	206	223	240	258	275	292	310	327
43	42	58	74	91	107	124	142	159	176	194	211	229	247	264	282	300	318	335
44	43	59	76	93	110	128	145	163	181	199	216	235	253	271	289	307	325	344
45	44	61	78	95	113	131	149	167	185	203	222	240	259	277	296	315	333	352
46	45	62	80	97	115	134	152	171	189	208	227	246	265	284	303	322	341	360
47	46	64	81	100	118	137	156	175	194	213	232	251	271	290	310	329	349	369
48	47	65	83	102	121	140	159	178	198	218	237	257	277	297	317	337	357	377
49	48	66	85	104	123	143	163	182	202	222	243	263	283	303	324	344	365	385
50	49	68	87	106	126	146	166	186	207	227	248	268	289	310	331	352	372	393
51	50	69	89	109	129	149	170	190	211	232	253	274	295	316	338	359	380	402
52	51	71	91	111	131	152	173	194	215	237	258	280	301	323	345	366	388	410
53	52	72	92	113	134	155	177	198	220	241	263	285	307	329	352	374	396	418
54	53	74	94	115	137	158	180	202	224	246	269	291	313	336	359	381	404	427
55	54	75	96	118	139	161	184	206	228	251	274	297	319	342	365	389	412	435
56	55	76	98	120	142	164	187	210	233	256	279	302	326	349	372	396	420	443
57	57	78	100	122	145	167	191	214	237	261	284	308	332	355	379	403	427	451
58	58	79	102	124	147	171	194	218	241	265	289	314	338	362	386	411	435	460
59	59	81	103	127	150	174	198	222	246	270	295	319	344	369	393	418	443	468
60	60	82	105	129	153	177	201	225	250	275	300	325	350	375	400	426	451	476
	p=0,01																	
41	23	36	49	63	77	91	106	121	136	151	166	181	196	211	227	242	258	273
42	23	37	50	65	79	94	109	124	139	155	170	186	201	217	233	249	265	280
43	24	38	52	66	81	96	112	127	143	159	175	190	207	223	239	255	271	288
44	25	39	53	68	83	99	115	130	146	163	179	195	212	228	245	262	278	295
45	25	40	54	70	85	101	117	134	150	167	183	200	217	234	251	268	285	303
46	26	41	56	71	87	104	120	137	154	171	188	205	222	240	257	275	292	310
47	27	42	57	73	90	106	123	140	157	175	192	210	228	245	263	281	299	317
48	27	43	58	75	92	109	126	143	161	179	197	215	233	251	269	288	306	325
49	28	44	60	77	94	111	129	147	165	183	201	220	238	257	276	294	313	332
50	29	45	61	78	96	114	132	150	168	187	206	225	244	263	282	301	320	339
51	29	46	63	80	98	116	135	153	172	191	210	229	249	268	288	307	327	347
52	30	47	64	82	100	119	137	157	176	195	215	234	254	274	294	314	334	354
53	31	48	65	83	102	121	140	160	179	199	219	239	259	280	300	320	341	361
54	31	49	67	85	104	114	143	163	183	203	224	244	265	285	306	327	348	369
55	32	50	68	87	106	126	146	166	187	207	228	249	270	291	312	333	355	376
56	33	51	69	89	108	129	149	177	190	211	233	254	275	297	318	340	362	384
57	33	52	71	90	111	131	152	173	194	215	237	259	281	302	324	347	369	391
58	34	53	72	92	113	133	155	176	198	220	242	264	286	308	331	353	376	398
59	34	54	73	94	115	136	158	179	201	224	246	268	291	314	337	360	383	406
60	35	55	75	96	117	138	160	183	205	228	250	273	296	320	343	366	390	413

n_1 / n_2	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	p=0,05																		
41	336	353	370	387	404	421	438	456	473	490	507	524	541	559	576	593	610	628	645
42	345	362	380	397	415	432	450	467	485	503	520	538	556	573	591	609	626	644	662
43	353	371	389	407	425	443	461	479	497	515	533	552	570	588	606	624	642	660	679
44	362	380	399	417	436	454	473	491	510	528	547	565	584	602	621	640	658	677	695
45	371	390	408	427	446	465	484	503	522	541	560	579	598	617	636	655	674	693	712
46	380	399	418	437	457	476	495	515	534	554	573	593	612	631	651	670	690	709	729
47	388	408	428	447	467	487	507	527	547	566	586	606	626	646	666	686	706	726	746
48	397	417	437	458	478	498	518	539	559	579	600	620	640	661	681	701	722	742	763
49	406	426	447	466	488	509	530	550	571	592	613	634	654	675	696	717	738	759	780
50	414	435	457	478	499	520	541	562	583	605	626	647	669	690	711	732	754	775	796
51	423	445	466	488	509	531	553	574	596	618	639	661	683	704	726	748	770	791	813
52	432	454	476	498	520	542	564	586	608	630	652	675	697	719	741	763	786	808	830
53	441	463	485	508	530	553	575	598	620	643	666	688	711	734	756	779	802	824	847
54	449	472	495	518	541	564	587	610	633	656	679	702	725	748	771	794	818	841	864
55	458	481	505	528	551	575	598	622	645	669	692	716	739	763	786	810	834	857	881
56	467	491	514	538	562	586	610	634	657	681	705	729	753	777	801	825	850	874	898
57	476	500	524	548	572	597	621	645	670	694	719	743	768	792	816	841	865	890	915
58	484	509	534	558	583	608	633	657	682	707	732	757	782	807	832	856	881	906	931
59	493	518	543	568	594	619	644	669	694	720	745	770	796	821	847	872	897	923	948
60	502	527	553	578	604	630	655	681	707	733	758	784	810	836	862	888	913	939	965
	p=0,01																		
41	289	304	320	336	351	367	383	398	414	430	446	462	477	493	509	525	541	557	573
42	296	312	328	345	361	377	393	409	425	442	458	474	490	507	523	539	556	572	588
43	304	321	337	354	370	387	403	420	437	453	470	487	503	520	537	553	570	587	604
44	312	329	346	363	380	397	414	431	448	465	482	499	516	533	550	568	585	602	619
45	320	337	354	372	389	407	424	441	459	476	494	511	529	547	564	582	599	617	635
46	328	345	363	381	399	416	434	452	470	488	506	524	542	560	578	596	614	632	650
47	335	353	372	390	408	426	445	463	481	500	518	536	555	573	592	610	629	647	666
48	343	362	380	399	418	436	455	474	492	511	530	549	568	587	606	625	643	662	681
49	351	370	389	408	427	446	465	484	504	523	542	561	581	600	619	639	658	678	697
50	359	378	398	417	437	456	476	495	515	535	554	574	594	613	633	653	673	693	713
51	366	386	406	526	446	466	486	506	526	546	566	587	607	627	647	667	688	708	728
52	374	395	415	435	456	476	496	517	537	558	578	599	620	640	661	682	702	723	744
53	382	403	423	444	465	486	507	528	549	570	591	612	633	654	675	696	717	738	759
54	390	411	432	453	475	496	517	538	560	581	603	624	646	667	689	710	732	753	775
55	398	419	441	462	484	506	527	549	571	593	615	637	659	680	702	724	746	768	790
56	405	427	449	471	494	516	538	560	582	605	627	649	671	694	716	738	761	784	806
57	413	436	458	581	503	526	548	571	593	616	639	662	684	707	730	753	776	799	822
58	421	444	467	490	513	536	559	582	605	628	651	674	697	721	744	767	790	814	837
59	429	452	475	499	522	545	569	592	616	640	663	687	710	734	758	781	805	829	853
60	437	460	484	508	532	555	579	603	627	651	675	699	723	747	772	796	820	844	868

n_1 / n_2	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
	p=0,05																			
41	662																			
42	679	697																		
43	697	715	733																	
44	714	733	751	770																
45	731	750	769	789	808															
46	749	768	788	807	827	846														
47	766	786	806	826	846	866	886													
48	783	804	824	845	865	886	906	927												
49	800	821	842	863	884	905	926	947	968											
50	818	839	861	882	903	925	946	968	989	1010										
51	835	857	879	901	922	944	966	988	1010	1032	1054									
52	852	875	897	919	942	964	986	1009	1031	1053	1076	1098								
53	870	893	915	938	961	984	1006	1029	1052	1075	1098	1120	1143							
54	887	910	934	957	980	1003	1026	1050	1073	1096	1119	1143	1166	1189						
55	904	928	952	975	999	1023	1046	1070	1094	1118	1142	1165	1189	1213	1236					
56	922	946	970	994	1018	1042	1067	1091	1115	1139	1163	1187	1212	1236	1260	1284				
57	939	964	988	1013	1037	1062	1087	1111	1136	1161	1185	1210	1235	1259	1284	1309	1333			
58	956	981	1007	1032	1057	1082	1107	1132	1157	1182	1207	1232	1257	1283	1308	1333	1358	1383		
59	974	999	1025	1050	1076	1101	1127	1152	1178	1204	1229	1255	1280	1306	1331	1357	1383	1408	1434	
60	991	1017	1043	1069	1095	1121	1147	1173	1199	1225	1251	1277	1303	1329	1355	1381	1407	1433	1460	1486
	p=0,01																			
41	589																			
42	605	621																		
43	621	637	654																	
44	636	654	671	688																
45	652	670	688	706	723															
46	668	687	705	723	741	759														
47	684	703	722	740	759	777	796													
48	700	719	738	757	776	795	814	834												
49	716	736	755	775	794	814	833	853	872											
50	732	752	772	792	812	832	852	872	892	912										
51	748	769	789	809	830	850	870	891	911	932	952									
52	764	785	806	827	847	868	889	910	931	951	972	993								
53	780	802	823	844	865	886	908	929	958	971	993	1014	1035							
54	796	818	840	861	883	905	926	948	970	991	1013	1035	1057	1078						
55	812	834	857	879	901	923	945	967	989	1011	1033	1056	1078	1100	1122					
56	828	851	873	896	919	941	964	986	1009	1031	1054	1077	1099	1122	1145	1167				
57	844	867	890	913	936	959	982	1005	1028	1051	1074	1098	1121	1144	1167	1190	1213			
58	861	884	907	931	954	978	1001	1024	1048	1071	1095	1118	1142	1165	1189	1213	1236	1260		
59	877	900	924	948	972	996	1020	1044	1068	1091	1115	1139	1163	1187	1211	1235	1259	1283	1307	
60	893	917	941	965	990	1014	1038	1062	1087	1111	1136	1160	1185	1209	1234	1258	1283	1307	1331	1356

Таблица III

Критические значения критерия H Крускала-Уоллиса для разных сочетаний n_1, n_2 и n_3

Различия между тремя выборками можно считать достоверными на указанном в таблице уровне значимости, если $H_{эмп}$ достигает соответствующего критического значения или превышает его (по Greene J., D'Olivera M., 1989).

Объемы выборок n_1	n_2	n_3	H	p
2	1	1	2,7000	0,500
2	2	1	3,6000	0,200
2	2	2	4,5714	0,067
3	1	1	3,2000	0,300
3	2	1	4,2857	0,100
			3,8571	0,133
3	2	2	5,3572	0,029
			4,7143	0,048
			4,5000	0,067
			4,4643	0,105
3	3	1	5,1429	0,043
			4,5714	0,100
			4,0000	0,129
3	3	2	6,2500	0,011
			5,3611	0,032
			5,1389	0,061
			4,5556	0,100
			4,2500	0,121
3	3	3	7,2000	0,004
			6,4889	0,011
			5,6889	0,029
			5,6000	0,050
			5,0667	0,086
			4,6222	0,100
4	1	1	3,5714	0,200
4	2	1	4,8214	0,057
			4,5000	0,076
			4,0179	0,114
4	2	2	6,0000	0,014
			5,3333	0,033
			5,1250	0,052
			4,4583	0,100
			4,1667	0,105
4	3	1	5,8333	0,021
			5,2083	0,050
			5,0000	0,057
			4,0556	0,093
			3,8889	0,129
4	3	2	6,4444	0,008
			6,3000	0,011
			5,4444	0,046
			5,4000	0,051
			4,5111	0,098
			4,4444	0,102
4	3	3	6,7455	0,010
			6,7091	0,013
			5,7909	0,046
			5,7273	0,050
			4,7091	0,092
			4,7000	0,101
4	4	1	6,6667	0,010
			6,1667	0,022
			4,9667	0,048
			4,8667	0,054
			4,1667	0,082
			4,0667	0,102
4	4	2	7,0364	0,006
			6,8727	0,011
			5,4545	0,046
			5,2364	0,052
			4,5545	0,098
			4,4455	0,103
4	4	3	7,1439	0,010
			7,1364	0,011
			5,5985	0,049
			5,5758	0,051
			4,5455	0,099
			4,4773	0,102
4	4	4	7,6538	0,008
			7,5385	0,011
			5,6923	0,049
			5,6538	0,054
			4,6539	0,097
			4,5001	0,104
5	1	1	3,8571	0,143
5	2	1	5,2500	0,036
			5,0000	0,048
			4,4500	0,071
			4,2000	0,095
			4,0500	0,119
5	2	2	6,5333	0,008
			6,1333	0,013
			5,1600	0,034
			5,0400	0,056
			4,3733	0,090
			4,2933	0,122
5	3	1	6,4000	0,012
			4,9600	0,048
			4,8711	0,052
			4,0178	0,095
			3,8400	0,123
5	3	2	6,9091	0,009
			6,8218	0,010
			5,2509	0,049
			5,1055	0,052
			4,6509	0,091
			4,4945	0,101
5	3	3	7,0788	0,009
			6,9818	0,011
			5,6485	0,049
			5,5152	0,051
			4,5333	0,097
			4,4121	0,109
5	4	1	6,9545	0,008
			6,8400	0,011
			4,9855	0,044
			4,8600	0,056
			3,9873	0,098
			3,9600	0,102
5	4	2	7,2045	0,009
			7,1182	0,010
			5,2727	0,049
			5,2682	0,050
			4,5409	0,098
			4,5182	0,101
5	4	3	7,4449	0,010
			7,3949	0,011
			5,6564	0,049
			5,6308	0,050
			4,5487	0,099
			4,5231	0,103
5	4	4	7,7604	0,009
			7,7440	0,011
			5,6571	0,049
			5,6176	0,050
			4,6187	0,100
			4,5527	0,102
5	5	1	7,3091	0,009
			6,8364	0,011
			5,1273	0,046
			4,9091	0,053
			4,1091	0,086
			4,0364	0,105
5	5	2	7,3385	0,010
			7,2692	0,010
			5,3385	0,047
			5,2462	0,051
			4,6231	0,097
			4,5077	0,100
5	5	3	7,5780	0,010
			7,5429	0,010
			5,7055	0,046
			5,6264	0,051
			4,5451	0,100
			4,5363	0,102
5	5	4	7,8229	0,010
			7,7914	0,010
			5,6657	0,049
			5,6429	0,050
			4,5229	0,099
			4,5200	0,101
5	5	5	8,0000	0,009
			7,9800	0,010
			5,7800	0,049
			5,6600	0,051
			4,5600	0,100
			4,5000	0,102

Таблица IV

Критические значения критерия тенденций S Джонкира для количества групп (c) от трех до шести ($3 \leq c \leq 6$) и количества испытуемых в каждой группе от двух до десяти ($2 \leq n \leq 10$)

Тенденция является достоверной, если $S_{эмп}$ достигает $S_{0,05}$ или превышает его ($\rho \leq 0,05$); и тем более достоверной, если $S_{эмп}$ достигает $S_{0,01}$ или превышает его ($\rho \leq 0,01$) (по Greene J., D'Olivera M., 1989).

c	N: 2	3	4	5	6	7	8	9	10
					ρ=0,05				
3	10	17	24	33	42	53	64	76	88
4	14	26	38	51	66	82	100	118	138
5	20	34	51	71	92	115	140	166	194
6	26	44	67	93	121	151	184	219	256
					ρ=0,01				
3	-	23	32	45	59	74	90	106	124
4	20	34	50	71	92	115	140	167	195
5	26	48	72	99	129	162	197	234	274
6	34	62	94	130	170	213	260	309	361

Таблица V

Критические значения критерия знаков G для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ (по Оуэну Д.Б., 1966)

Преобладание "типичного" сдвига является достоверным, если $G_{эмп}$ ниже или равен $G_{0,05}$, и тем более достоверным, если $G_{эмп}$ ниже или равен $G_{0,01}$.

n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01
5	0	-	27	8	7	49	18	15	92	37	34
6	0	-	28	8	7	50	18	16	94	38	35
7	0	0	29	9	7	52	19	17	96	39	36
8	1	0	30	10	8	54	20	18	98	40	37
9	1	0	31	10	8	56	21	18	100	41	37
10	1	0	32	10	8	58	22	19	110	45	42
11	2	1	33	11	9	60	23	20	120	50	46
12	2	1	34	11	9	62	24	21	130	55	51
13	3	1	35	12	10	64	24	22	140	59	55
14	3	2	36	12	10	66	25	23	150	64	60
15	3	2	37	13	10	68	26	23	160	69	64
16	4	2	38	13	11	70	27	24	170	73	69
17	4	3	39	13	11	72	28	25	180	78	73
18	5	3	40	14	12	74	29	26	190	83	78
19	5	4	41	14	12	76	30	27	200	87	83
20	5	4	42	15	13	78	31	28	220	97	92
21	6	4	43	15	13	80	32	29	240	106	101
22	6	5	44	16	13	82	33	30	260	116	110
23	7	5	45	16	14	84	33	30	280	125	120
24	7	5	46	16	14	86	34	31	300	135	129
25	7	6	47	17	15	88	35	32			
26	8	6	48	17	15	90	36	33			

Таблица VI

Критические значения критерия T Вилкоксона для уровней статистической значимости $\rho \leq 0{,}05$ и $\rho \leq 0{,}01$

"Типичный" сдвиг является достоверно преобладающим по интенсивности, если $T_{эмп}$ ниже или равен $T_{0,05}$, и тем более достоверно преобладающим, если $T_{эмп}$ ниже или равен $T_{0,01}$ (по Wilcoxon F. et al., 1963).

n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01
5	0	–	28	130	101
6	2	–	29	140	110
7	3	0	30	151	120
8	5	1	31	163	130
9	8	3	32	175	140
10	10	5	33	187	151
11	13	7	34	200	162
12	17	9	35	213	173
13	21	12	36	227	185
14	25	15	37	241	198
15	30	19	38	256	211
16	35	23	39	271	224
17	41	27	40	286	238
18	47	32	41	302	252
19	53	37	42	319	266
20	60	43	43	336	281
21	67	49	44	353	296
22	75	55	45	371	312
23	83	62	46	389	328
24	91	69	47	407	345
25	100	76	48	426	362
26	110	84	49	446	379
27	119	92	50	466	397

Таблица VII-A

Критические значения критерия χ^2_r Фридмана для количества условий $c=3$ и количества испытуемых от двух до девяти ($2 \leq n \leq 9$)

Различия между условиями можно считать достоверными на указанном в таблице уровне значимости, если $\chi^2_{r\ эмп}$ достигает соответствующего критического значения или превышает его (по Greene J., D'Olivera M., 1989).

n=2		n=3		n=4		n=5	
χ^2_r	ρ	χ^2_r	ρ	χ^2_r	ρ	χ^2_r	ρ
0	1,000	0,000	1,000	0,0	1,000	0,0	1,000
1	0,833	0,667	0,944	0,5	0,931	0,4	0,954
3	0,500	2,000	0,528	1,5	0,653	1,2	0,691
4	0,167	2,667	0,361	2,0	0,431	1,6	0,522
		4,667	0,194	3,5	0,273	2,8	0,367
		6,000	0,028	4,5	0,125	3,6	0,182
				6,0	0,069	4,8	0,124
				6,5	0,042	5,2	0,093
				8,0	0,0046	6,4	0,039
						7,6	0,024
						8,4	0,0085
						10,0	0,00077

n=6		n=7		n=8		n=9	
χ^2_r	ρ	χ^2_r	ρ	χ^2_r	ρ	χ^2_r	ρ
0,00	1,000	0,000	1,000	0,00	1,000	0,000	1,000
0,33	0,956	0,286	0,964	0,25	0,967	0,222	0,971
1,00	0,740	0,857	0,768	0,75	0,794	0,667	0,814
1,33	0,570	1,143	0,620	1,00	0,654	0,889	0,865
2,33	0,430	2,000	0,486	1,75	0,531	1,556	0,569
3,00	0,252	2,571	0,305	2,25	0,355	2,000	0,398
4,00	0,184	3,429	0,237	3,00	0,285	2,667	0,328
4,33	0,142	3,714	0,192	3,25	0,236	2,889	0,278
5,33	0,072	4,571	0,112	4,00	0,149	3,556	0,187
6,33	0,052	5,429	0,085	4,75	0,120	4,222	0,154
7,00	0,029	6,000	0,052	5,25	0,079	4,667	0,107
8,33	0,012	7,143	0,027	6,25	0,047	5,556	0,069
9,00	0,0081	7,714	0,021	6,75	0,038	6,000	0,057
9,33	0,0055	8,000	0,016	7,00	0,030	6,222	0,048
10,33	0,0017	8,857	0,0084	7,75	0,018	6,889	0,031
12,00	0,00013	10,286	0,0036	9,00	0,0099	8,000	0,019
		10,571	0,0027	9,25	0,0080	8,222	0,016
		11,143	0,0012	9,75	0,0048	8,667	0,010
		12,286	0,00032	10,75	0,0024	9,556	0,0060
		14,000	0,000021	12,00	0,0011	10,667	0,0035
				12,25	0,00086	10,889	0,0029
				13,00	0,00026	11,556	0,0013
				14,25	0,000061	12,667	0,00066
				16,00	0,0000036	13,556	0,00035
						14,000	0,00020
						14,222	0,000097
						14,889	0,000054
						16,222	0,000011
						18,000	0,0000006

Таблица VII-Б

Критические значения критерия χ^2_r Фридмана для количества условий c=4, $2 \leq n \leq 4$

Различия между условиями можно считать достоверными на указанном в таблице уровне значимости, если $\chi^2_{r\ эмп}$ достигает соответствующего критического значения или превышает его (по Greene J., D'Olivera M., 1989).

n=2		n=3		n=4			
χ^2_r	p	χ^2_r	p	χ^2_r	p	χ^2_r	p
0,0	1,000	0,0	1,000	0,0	1,000	5,7	0,141
0,6	0,958	0,6	0,958	0,3	0,992	6,0	0,105
1,2	0,834	1,0	0,910	0,6	0,928	6,3	0,094
1,8	0,792	1,8	0,727	0,9	0,900	6,6	0,077
2,4	0,625	2,2	0,608	1,2	0,800	6,9	0,068
3,0	0,542	2,6	0,524	1,5	0,754	7,2	0,054
3,6	0,458	3,4	0,446	1,8	0,677	7,5	0,052
4,2	0,375	3,8	0,342	2,1	0,649	7,8	0,036
4,8	0,208	4,2	0,300	2,4	0,524	8,1	0,033
5,4	0,167	5,0	0,207	2,7	0,508	8,4	0,019
6,0	0,042	5,4	0,175	3,0	0,432	8,7	0,014
		5,8	0,148	3,3	0,389	9,3	0,012
		6,6	0,075	3,6	0,355	9,6	0,0069
		7,0	0,054	3,9	0,324	9,9	0,0062
		7,4	0,033	4,5	0,242	10,2	0,0027
		8,2	0,017	4,8	0,200	10,8	0,0016
		9,0	0,0017	5,1	0,190	11,1	0,00094
				5,4	0,158	12,0	0,000072

Таблица VIII

Критические значения критерия тенденций L Пейджа для количества условий от трех до шести ($3 \leq c \leq 6$) и количества испытуемых от двух до двенадцати ($2 \leq n \leq 12$)

Тенденция является достоверной, если $L_{эмп}$ достигает или превышает $L_{0,05}$, и тем более достоверной, если $L_{эмп}$ достигает или превышает $L_{0,01}$ (по Greene J., D'Olivera M., 1989).

n	c (количество условий)				
	3	4	5	6	p
2	–	–	109	178	0,001
	–	60	106	173	0,01
	28	58	103	166	0,05
3	–	89	160	260	0,001
	42	87	155	252	0,01
	41	84	150	244	0,05
4	56	117	210	341	0,001
	55	114	204	331	0,01
	54	111	197	321	0,05
5	70	145	259	420	0,001
	68	141	251	409	0,01
	66	137	244	397	0,05
6	83	172	307	499	0,001
	81	167	299	486	0,01
	79	163	291	474	0,05
7	96	198	355	577	0,001
	93	193	346	563	0,01
	91	189	338	550	0,05
8	109	225	403	655	0,001
	106	220	393	640	0,01
	104	214	384	625	0,05
9	121	252	451	733	0,001
	119	246	441	717	0,01
	116	240	431	701	0,05
10	134	278	499	811	0,001
	131	272	487	793	0,01
	128	266	477	777	0,05
11	147	305	546	888	0,001
	144	298	534	869	0,01
	141	292	523	852	0,05
12	160	331	593	965	0,001
	156	324	581	946	0,01
	153	317	570	928	0,05

Таблица IX

Критические значения критерия χ^2 для уровней статистической значимости $p \leq 0{,}05$ и $p \leq 0{,}01$ при разном числе степеней свободы ν

Различия между двумя распределениями могут считаться достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,05}$, и тем более достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,01}$ (по Большеву Л.Н., Смирнову Н.В., 1983).

ν	p 0,05	p 0,01	ν	p 0,05	p 0,01	ν	p 0,05	p 0,01
1	3,841	6,635	35	49,802	57,342	69	89,391	99,227
2	5,991	9,210	36	50,998	58,619	70	90,631	100,425
3	7,815	11,345	37	52,192	59,892	71	91,670	101,621
4	9,488	13,277	38	53,384	61,162	72	92,808	102,816
5	11,070	15,086	39	54,572	62,428	73	93,945	104,010
6	12,592	16,812	40	55,758	63,691	74	95,081	105,202
7	14,067	18,475	41	56,942	64,950	75	96,217	106,393
8	15,507	20,090	42	58,124	66,206	76	97,351	107,582
9	16,919	21,666	43	59,304	67,459	77	98,484	108,771
10	18,307	23,209	44	60,481	68,709	78	99,617	109,958
11	19,675	24,725	45	61,656	69,957	79	100,749	111,144
12	21,026	26,217	46	62,830	71,201	80	101,879	112,329
13	22,362	27,688	47	64,001	72,443	81	103,010	113,512
14	23,685	29,141	48	65,171	73,683	82	104,139	114,695
15	24,996	30,578	49	66,339	74,919	83	105,267	115,676
16	26,296	32,000	50	67,505	76,154	84	106,395	117,057
17	27,587	33,409	51	68,669	77,386	85	107,522	118,236
18	28,869	34,805	52	69,832	78,616	86	108,648	119,414
19	30,144	36,191	53	70,993	79,843	87	109,773	120,591
20	31,410	37,566	54	72,153	81,069	88	110,898	121,767
21	32,671	38,932	55	73,311	82,292	89	112,022	122,942
22	33,924	40,289	56	74,468	83,513	90	113,145	124,116
23	35,172	41,638	57	75,624	84,733	91	114,268	125,289
24	36,415	42,980	58	76,778	85,950	92	115,390	126,462
25	37,652	44,314	59	77,931	87,166	93	116,511	127,633
26	38,885	45,642	60	79,082	88,379	94	117,632	128,803
27	40,113	46,963	61	80,232	89,591	95	118,752	129,973
28	41,337	48,278	62	81,381	90,802	96	119,871	131,141
29	42,557	49,588	63	82,529	92,010	97	120,990	132,309
30	43,773	50,892	64	83,675	93,217	98	122,108	133,476
31	44,985	52,191	65	84,821	94,422	99	123,225	134,642
32	46,194	53,486	66	85,965	95,626	100	124,342	135,807
33	47,400	54,776	67	87,108	96,828			
34	48,602	56,061	68	88,250	98,028			

Таблица X

Критические значения d_{max}, соответствующие уровням статистической значимости $p \leq 0{,}05$ и $p \leq 0{,}01$ при сопоставлении эмпирического распределения с теоретическим

Различия между распределениями могут считаться достоверными, если абсолютная величина максимальной разности $d_{эмп}$ достигает или превышает $d_{0,05}$, и тем более достоверными, если $d_{эмп}$ достигает или превышает $d_{0,01}$ (по Ван дер Вардену Б.Л., 1960).

n	Максимальный модуль разности накопленных частостей d_{max}		n	Максимальный модуль разности накопленных частостей d_{max}	
	p=0,05	p=0,01		p=0,05	p=0,01
5	0,6074	0,7279	50	0,1921	0,2302
10	0,4295	0,5147	60	0,1753	0,2101
15	0,3507	0,4202	70	0,1623	0,1945
20	0,3037	0,3639	80	0,1518	0,1820
25	0,2716	0,3255	90	0,1432	
30	0,2480	0,2972	100	0,1358	
40	0,2147	0,2574	>100	$1{,}36/\sqrt{n}$	$1{,}63/\sqrt{n}$

Таблица XI

Критерий λ Колмогорова-Смирнова для сопоставления эмпирического распределения с теоретическим (при $n>50$) или двух эмпирических распределений между собой (при $n>50$): уровни статистической значимости разных значений $\lambda_{эмп}$

По полученному значению $\lambda_{эмп}$ определяется уровень значимости различий между двумя распределениями (по Митропольскому А.К., 1971).

λ	λ, последний десятичный знак									
	0	1	2	3	4	5	6	7	8	9
	p - десятичные знаки («0,» опущен)									
0,3	99999	99998	99995	99991	99983	99970	99949	99917	99872	99807
0,4	99719	99603	99452	99262	99027	98741	98400	97998	97532	96998
0,5	96394	95719	94969	94147	93250	92282	91242	90134	88960	87724
0,6	86428	85077	83678	82225	80732	79201	77636	76042	74422	72781
0,7	71124	69453	67774	66089	64402	62717	61036	59363	57700	56050
0,8	54414	52796	51197	49619	48063	46532	45026	43545	42093	40668
0,9	39273	37907	36571	35266	33992	32748	31536	30356	29206	28087
1,0	27000	25943	24917	23922	22957	22021	21114	20236	19387	18566
1,1	17772	17005	16264	15550	14861	14196	13556	12939	12345	11774
1,2	11225	10697	10190	09703	09235	08787	08357	07944	07550	07171
1,3	06809	06463	06132	05815	05513	05224	04949	04686	04435	04196
1,4	03968	03751	03545	03348	03162	02984	02815	02655	02503	02359
1,5	02222	02092	01969	01852	01742	01638	01539	01446	01357	01274
1,6	01195	01121	01051	00985	00922	00864	00808	00756	00707	00661
1,7	00618	00577	00539	00503	00469	00438	00408	00380	00354	00330
1,8	00307	00285	00265	00247	00229	00213	00198	00186	00170	00158
1,9	00146	00136	00126	00116	00108	00100	00092	00085	00079	00073
2,0	00067	00062	00057	00053	00048	00045	00041	00038	00035	00032
2,1	00030	00027	00025	00023	00021	00019	00018	00016	00015	00014
2,2	00013	00011	00010	00010	00009	00008	00007	00007	00006	00006
2,3	00005	00005	00004	00004	00004	00003	00003	00003	00002	00002
2,4	00002	00002	00002	00001	00001	00001	00001	00001	00001	00001

Таблица XII

Величины угла φ (в радианах) для разных процентных долей: φ=2·arcsin√P (по Урбаху В.Ю., 1964)

%доля	%, последний десятичный знак: 0	1	2	3	4	5	6	7	8	9
	Значения φ=2·arcsin√P									
0,0	0,000	0,020	0,028	0,035	0,040	0,045	0,049	0,053	0,057	0,060
0,1	0,063	0,066	0,069	0,072	0,075	0,077	0,080	0,082	0,085	0,087
0,2	0,089	0,092	0,094	0,096	0,098	0,100	0,102	0,104	0,106	0,108
0,3	0,110	0,111	0,113	0,115	0,117	0,118	0,120	0,122	0,123	0,125
0,4	0,127	0,128	0,130	0,131	0,133	0,134	0,136	0,137	0,139	0,140
0,5	0,142	0,143	0,144	0,146	0,147	0,148	0,150	0,151	0,153	0,154
0,6	0,155	0,156	0,158	0,159	0,160	0,161	0,163	0,164	0,165	0,166
0,7	0,168	0,169	0,170	0,171	0,172	0,173	0,175	0,176	0,177	0,178
0,8	0,179	0,180	0,182	0,183	0,184	0,185	0,186	0,187	0,188	0,189
0,9	0,190	0,191	0,192	0,193	0,194	0,195	0,196	0,197	0,198	0,199
1	0,200	0,210	0,220	0,229	0,237	0,246	0,254	0,262	0,269	0,277
2	0,284	0,291	0,298	0,304	0,311	0,318	0,324	0,330	0,336	0,342
3	0,348	0,354	0,360	0,365	0,371	0,376	0,382	0,387	0,392	0,398
4	0,403	0,408	0,413	0,418	0,423	0,428	0,432	0,437	0,442	0,446
5	0,451	0,456	0,460	0,465	0,469	0,473	0,478	0,482	0,486	0,491
6	0,495	0,499	0,503	0,507	0,512	0,516	0,520	0,524	0,528	0,532
7	0,536	0,539	0,543	0,547	0,551	0,555	0,559	0,562	0,566	0,570
8	0,574	0,577	0,581	0,584	0,588	0,592	0,595	0,599	0,602	0,606
9	0,609	0,613	0,616	0,620	0,623	0,627	0,630	0,633	0,637	0,640
10	0,644	0,647	0,650	0,653	0,657	0,660	0,663	0,666	0,670	0,673
11	0,676	0,679	0,682	0,686	0,689	0,692	0,695	0,698	0,701	0,704
12	0,707	0,711	0,714	0,717	0,720	0,723	0,726	0,729	0,732	0,735
13	0,738	0,741	0,744	0,747	0,750	0,752	0,755	0,758	0,761	0,764
14	0,767	0,770	0,773	0,776	0,778	0,781	0,784	0,787	0,790	0,793
15	0,795	0,798	0,801	0,804	0,807	0,809	0,812	0,815	0,818	0,820
16	0,823	0,826	0,828	0,831	0,834	0,837	0,839	0,842	0,845	0,847
17	0,850	0,853	0,855	0,858	0,861	0,863	0,866	0,868	0,871	0,874
18	0,876	0,879	0,881	0,884	0,887	0,889	0,892	0,894	0,897	0,900
19	0,902	0,905	0,907	0,910	0,912	0,915	0,917	0,920	0,922	0,925
20	0,927	0,930	0,932	0,935	0,937	0,940	0,942	0,945	0,947	0,950
21	0,952	0,955	0,957	0,959	0,962	0,964	0,967	0,969	0,972	0,974
22	0,976	0,979	0,981	0,984	0,986	0,988	0,991	0,993	0,996	0,998
23	1,000	1,003	1,005	1,007	1,010	1,012	1,015	1,017	1,019	1,022
24	1,024	1,026	1,029	1,031	1,033	1,036	1,038	1,040	1,043	1,045
25	1,047	1,050	1,052	1,054	1,056	1,059	1,061	1,063	1,066	1,068
26	1,070	1,072	1,075	1,077	1,079	1,082	1,084	1,086	1,088	1,091
27	1,093	1,095	1,097	1,100	1,102	1,104	1,106	1,109	1,111	1,113
28	1,115	1,117	1,120	1,122	1,124	1,126	1,129	1,131	1,133	1,135
29	1,137	1,140	1,142	1,144	1,146	1,148	1,151	1,153	1,155	1,157
30	1,159	1,161	1,164	1,166	1,168	1,170	1,172	1,174	1,177	1,179

Таблица XII. Продолжение

%доли	%, последний десятичный знак: 0	1	2	3	4	5	6	7	8	9
	Значения φ=2·arcsin√P									
31	1,182	1,183	1,185	1,187	1,190	1,192	1,194	1,196	1,198	1,200
32	1,203	1,205	1,207	1,209	1,211	1,213	1,215	1,217	1,220	1,222
33	1,224	1,226	1,228	1,230	1,232	1,234	1,237	1,239	1,241	1,243
34	1,245	1,247	1,249	1,251	1,254	1,256	1,258	1,260	1,262	1,264
35	1,266	1,268	1,270	1,272	1,274	1,277	1,279	1,281	1,283	1,285
36	1,287	1,289	1,291	1,293	1,295	1,297	1,299	1,302	1,304	1,306
37	1,308	1,310	1,312	1,314	1,316	1,318	1,320	1,322	1,324	1,326
38	1,328	1,330	1,333	1,335	1,337	1,339	1,341	1,343	1,345	1,347
39	1,349	1,351	1,353	1,355	1,357	1,359	1,361	1,363	1,365	1,367
40	1,369	1,371	1,374	1,376	1,378	1,380	1,382	1,384	1,386	1,388
41	1,390	1,392	1,394	1,396	1,398	1,400	1,402	1,404	1,406	1,408
42	1,410	1,412	1,414	1,416	1,418	1,420	1,422	1,424	1,426	1,428
43	1,430	1,432	1,434	1,436	1,438	1,440	1,442	1,444	1,446	1,448
44	1,451	1,453	1,455	1,457	1,459	1,461	1,463	1,465	1,467	1,469
45	1,471	1,473	1,475	1,477	1,479	1,481	1,483	1,485	1,487	1,489
46	1,491	1,493	1,495	1,497	1,499	1,501	1,503	1,505	1,507	1,509
47	1,511	1,513	1,515	1,517	1,519	1,521	1,523	1,525	1,527	1,529
48	1,531	1,533	1,535	1,537	1,539	1,541	1,543	1,545	1,547	1,549
49	1,551	1,553	1,555	1,557	1,559	1,561	1,563	1,565	1,567	1,569
50	1,571	1,573	1,575	1,577	1,579	1,581	1,583	1,585	1,587	1,589
51	1,591	1,593	1,595	1,597	1,599	1,601	1,603	1,605	1,607	1,609
52	1,611	1,613	1,615	1,617	1,619	1,621	1,623	1,625	1,627	1,629
53	1,631	1,633	1,635	1,637	1,639	1,641	1,643	1,645	1,647	1,649
54	1,651	1,653	1,655	1,657	1,659	1,661	1,663	1,665	1,667	1,669
55	1,671	1,673	1,675	1,677	1,679	1,681	1,683	1,685	1,687	1,689
56	1,691	1,693	1,695	1,697	1,699	1,701	1,703	1,705	1,707	1,709
57	1,711	1,713	1,715	1,717	1,719	1,721	1,723	1,725	1,727	1,729
58	1,731	1,734	1,736	1,738	1,740	1,742	1,744	1,746	1,748	1,750
59	1,752	1,754	1,756	1,758	1,760	1,762	1,764	1,766	1,768	1,770
60	1,772	1,774	1,776	1,778	1,780	1,782	1,784	1,786	1,789	1,791
61	1,793	1,795	1,797	1,799	1,801	1,803	1,805	1,807	1,809	1,811
62	1,813	1,815	1,817	1,819	1,821	1,823	1,826	1,828	1,830	1,832
63	1,834	1,836	1,838	1,840	1,842	1,844	1,846	1,848	1,850	1,853
64	1,855	1,857	1,859	1,861	1,863	1,865	1,867	1,869	1,871	1,873
65	1,875	1,878	1,880	1,882	1,884	1,886	1,888	1,890	1,892	1,894
66	1,897	1,899	1,901	1,903	1,905	1,907	1,909	1,911	1,913	1,916
67	1,918	1,920	1,922	1,924	1,926	1,928	1,930	1,933	1,935	1,937
68	1,939	1,941	1,943	1,946	1,948	1,950	1,952	1,954	1,956	1,958
69	1,961	1,963	1,965	1,967	1,969	1,971	1,974	1,976	1,978	1,980
70	1,982	1,984	1,987	1,989	1,991	1,993	1,995	1,998	2,000	2,002
71	2,004	2,006	2,009	2,011	2,013	2,015	2,018	2,020	2,022	2,024
72	2,026	2,029	2,031	2,033	2,035	2,038	2,040	2,042	2,044	2,047
73	2,049	2,051	2,053	2,056	2,058	2,060	2,062	2,065	2,067	2,069
74	2,071	2,074	2,076	2,078	2,081	2,083	2,085	2,087	2,090	2,092
75	2,094	2,097	2,099	2,101	2,104	2,106	2,108	2,111	2,113	2,115
76	2,118	2,120	2,122	2,125	2,127	2,129	2,132	2,134	2,136	2,139
77	2,141	2,144	2,146	2,148	2,151	2,153	2,156	2,158	2,160	2,163
78	2,165	2,168	2,170	2,172	2,175	2,177	2,180	2,182	2,185	2,187
79	2,190	2,192	2,194	2,197	2,199	2,202	2,204	2,207	2,209	2,212
80	2,214	2,217	2,219	2,222	2,224	2,227	2,229	2,231	2,234	2,237

Таблица XII. Окончание

%-доля	% последний десятичный знак: 0	1	2	3	4	5	6	7	8	9
	Значения φ=2·arcsin√P									
81	2,240	2,242	2,245	2,247	2,250	2,252	2,255	2,258	2,260	2,263
82	2,265	2,268	2,271	2,273	2,276	2,278	2,281	2,284	2,286	2,289
83	2,292	2,294	2,297	2,300	2,302	2,305	2,308	2,310	2,313	2,316
84	2,319	2,321	2,324	2,327	2,330	2,332	2,335	2,338	2,341	2,343
85	2,346	2,349	2,352	2,355	2,357	2,360	2,363	2,366	2,369	2,372
86	2,375	2,377	2,380	2,383	2,386	2,389	2,392	2,395	2,398	2,401
87	2,404	2,407	2,410	2,413	2,416	2,419	2,422	2,425	2,428	2,431
88	2,434	2,437	2,440	2,443	2,447	2,450	2,453	2,456	2,459	2,462
89	2,465	2,469	2,472	2,475	2,478	2,482	2,485	2,488	2,491	2,495
90	2,498	2,501	2,505	2,508	2,512	2,515	2,518	2,522	2,525	2,529
91	2,532	2,536	2,539	2,543	2,546	2,550	2,554	2,557	2,561	2,564
92	2,568	2,572	2,575	2,579	2,583	2,587	2,591	2,594	2,598	2,602
93	2,606	2,610	2,614	2,618	2,622	2,626	2,630	2,634	2,638	2,642
94	2,647	2,651	2,655	2,659	2,664	2,668	2,673	2,677	2,681	2,686
95	2,691	2,295	2,700	2,705	2,709	2,714	2,719	2,724	2,729	2,734
96	2,739	2,744	2,749	2,754	2,760	2,765	2,771	2,776	2,782	2,788
97	2,793	2,799	2,805	2,811	2,818	2,824	2,830	2,837	2,844	2,851
98	2,858	2,865	2,872	2,880	2,888	2,896	2,904	2,913	2,922	2,931
99,0	2,941	2,942	2,943	2,944	2,945	2,946	2,948	2,949	2,950	2,951
99,1	2,952	2,953	2,954	2,955	2,956	2,957	2,958	2,959	2,960	2,961
99,2	2,963	2,964	2,965	2,966	2,967	2,968	2,969	2,971	2,972	2,973
99,3	2,974	2,975	2,976	2,978	2,979	2,980	2,981	2,983	2,984	2,985
99,4	2,987	2,988	2,989	2,990	2,992	2,993	2,995	2,996	2,997	2,999
99,5	3,000	3,002	3,003	3,004	3,006	3,007	3,009	3,010	3,012	3,013
99,6	3,015	3,017	3,018	3,020	3,022	3,023	3,025	3,027	3,028	3,030
99,7	3,032	3,034	3,036	3,038	3,040	3,041	3,044	3,046	3,048	3,050
99,8	3,052	3,054	3,057	3,059	3,062	3,064	3,067	3,069	3,072	3,075
99,9	3,078	3,082	3,085	3,089	3,093	3,097	3,101	3,107	3,113	3,122
100	3,142									

Таблица XIII

Уровни статистической значимости разных значений критерия φ* Фишера (по Гублеру Е.В., 1978)

По полученному значению $\varphi^*_{эмп}$ определяется уровень значимости различий процентных долей.

ρ равно или меньше	ρ равно или меньше (последний десятичный знак): 0	1	2	3	4	5	6	7	8	9
0,00	2,91	2,81	2,70	2,62	2,55	2,49	2,44	2,39	2,35	
0,01	2,31	2 28	2,25	2,22	2,19	2,16	2,14	2,11	2,09	2,07
0,02	2,05	2,03	2,01	1,99	1,97	1,96	1,94	1,92	1,91	1,89
0,03	1,88	1,86	1,85	1,84	1,82	1,81	1,80	1,79	1,77	1,76
0,04	1,75	1,74	1,73	1,72	1,71	1,70	1,68	1,67	1,66	1,65
0,05	1,64	1,64	1,63	1,62	1,61	1,60	1,59	1,58	1,57	1,56
0,06	1,56	1,55	1,54	1,53	1,52	1,52	1,51	1,50	1,49	1,48
0,07	1,48	1,47	1,46	1,46	1,45	1,44	1,43	1,43	1,42	1,41
0,08	1,41	1,40	1,39	1,39	1,38	1,37	1,37	1,36	1,36	1,35
0,09	1,34	1,34	1,33	1,32	1,32	1,31	1,31	1,30	1,30	1,29
0,10	1,29									

Таблица XIV

Критические значения биномиального критерия m при P=0,50, n≤300 (рассчитано по Оуэну Д.В., 1966)

Различия достоверны, если $m_{эмп}$ равен или больше $m_{0,05}$, и тем более достоверны, если $m_{эмп}$ равен или больше $m_{0,01}$

n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01	n	ρ 0,05	ρ 0,01
5	5	-	27	19	20	49	31	34	90	54	57
6	6	-	28	20	21	50	32	34	92	55	58
7	7	7	29	20	22	52	33	35	94	56	59
8	7	8	30	20	22	54	34	36	96	57	60
9	8	9	31	21	23	56	35	38	98	58	61
10	9	10	32	22	24	58	36	39	100	59	63
11	9	10	33	22	24	60	37	40	110	65	68
12	10	11	34	23	25	62	38	41	120	70	74
13	10	12	35	23	25	64	40	42	130	75	79
14	11	12	36	24	26	66	41	43	140	81	85
15	12	13	37	24	27	68	42	45	150	86	90
16	12	14	38	25	27	70	43	46	160	91	96
17	13	14	39	26	28	72	44	47	170	97	101
18	13	15	40	26	28	74	45	48	180	102	107
19	14	15	41	27	29	76	46	49	190	107	112
20	15	16	42	27	29	78	47	50	200	113	117
21	15	17	43	28	30	80	48	51	220	123	128
22	16	17	44	28	31	82	49	52	240	134	139
23	16	18	45	29	31	84	51	54	260	144	150
24	17	19	46	30	32	86	52	55	280	155	160
25	18	19	47	30	32	88	53	56	300	165	171
26	18	20	48	31	38						

Таблица XV

Критические значения биномиального критерия m при $P \leq 0,50$; $n \leq 50$ (по Руниону Р., 1982)

Различия достоверны, если $m_{эмп}$ равен или больше $m_{0,05}$, и тем более достоверны, если $m_{эмп}$ равен или больше $m_{0,01}$.

N \ P / Q	0,01 / 0,99	0,02 / 0,98	0,03 / 0,97	0,04 / 0,96	0,05 / 0,95	0,06 / 0,94	0,07 / 0,93	0,08 / 0,92	0,09 / 0,91	0,10 / 0,90	0,11 / 0,89	0,12 / 0,88	0,13 / 0,87	0,14 / 0,86	0,15 / 0,85	0,16 / 0,84	0,17 / 0,83
ρ=0,05																	
2	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3
4	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3
5	1	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3
6	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4
7	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4
8	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4
9	2	2	2	2	3	3	3	3	3	4	4	4	4	4	4	4	5
10	2	2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5
11	2	2	2	3	3	3	3	4	4	4	4	4	4	5	5	5	5
12	2	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5
13	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	6
14	2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	6	6
15	2	2	3	3	3	4	4	4	4	5	5	5	5	5	6	6	6
16	2	2	3	3	3	4	4	4	4	5	5	5	5	6	6	6	6
17	2	2	3	3	4	4	4	4	5	5	5	5	6	6	6	6	7
18	2	2	3	3	4	4	4	5	5	5	5	6	6	6	6	7	7
19	2	3	3	3	4	4	4	5	5	5	6	6	6	6	7	7	7
20	2	3	3	3	4	4	4	5	5	5	6	6	6	7	7	7	7
21	2	3	3	3	4	4	5	5	5	6	6	6	6	7	7	7	8
22	2	3	3	4	4	4	5	5	5	6	6	6	7	7	7	8	8
23	2	3	3	4	4	4	5	5	6	6	6	6	7	7	7	8	8
24	2	3	3	4	4	5	5	5	6	6	6	7	7	7	8	8	8
25	2	3	3	4	4	5	5	5	6	6	6	7	7	8	8	8	8
ρ=0,01																	
2	1	2	2	2	2	2	2	2	2	2	.	.	.	.	.	.	.
3	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
4	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4
5	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	4
6	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4
7	2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
8	2	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5
9	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	5
10	2	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6
11	2	3	3	3	4	4	4	4	5	5	5	5	5	6	6	6	6
12	2	3	3	4	4	4	4	5	5	5	5	5	6	6	6	6	6
13	2	3	3	4	4	4	5	5	5	5	5	6	6	6	6	7	7
14	2	3	3	4	4	4	5	5	5	5	6	6	6	6	7	7	7
15	2	3	3	4	4	5	5	5	5	6	6	6	6	7	7	7	7
16	3	3	4	4	4	5	5	5	6	6	6	6	7	7	7	7	8
17	3	3	4	4	4	5	5	5	6	6	6	7	7	7	7	8	8
18	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
19	3	3	4	4	5	5	5	6	6	6	7	7	7	8	8	8	8
20	3	3	4	4	5	5	6	6	6	7	7	7	8	8	8	8	9
21	3	3	4	4	5	5	6	6	6	7	7	7	8	8	8	9	9
22	3	3	4	5	5	5	6	6	7	7	7	8	8	8	9	9	9
23	3	4	4	5	5	6	6	6	7	7	7	8	8	9	9	9	9
24	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	10
25	3	4	4	5	5	6	6	7	7	7	8	8	9	9	9	10	10

Таблица XV. Продолжение

N \ P / Q	0,01 / 0,99	0,02 / 0,98	0,03 / 0,97	0,04 / 0,96	0,05 / 0,95	0,06 / 0,94	0,07 / 0,93	0,08 / 0,92	0,09 / 0,91	0,10 / 0,90	0,11 / 0,89	0,12 / 0,88	0,13 / 0,87	0,14 / 0,86	0,15 / 0,85	0,16 / 0,84	0,17 / 0,83
ρ=0,05																	
26	2	3	3	4	4	5	5	6	6	6	7	7	7	8	8	8	9
27	2	3	3	4	4	5	5	6	6	6	7	7	8	8	8	9	9
28	2	3	4	4	4	5	5	6	6	7	7	7	8	8	8	9	9
29	2	3	4	4	5	5	5	6	6	7	7	8	8	8	9	9	9
30	2	3	4	4	5	5	6	6	6	7	7	8	8	8	9	9	10
31	2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	10
32	2	3	4	4	5	5	6	6	7	7	8	8	8	9	9	10	10
33	2	3	4	4	5	5	6	6	7	7	8	8	9	9	9	10	10
34	2	3	4	4	5	6	6	7	7	7	8	8	9	9	10	10	11
35	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	11
36	3	3	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11
37	3	3	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11
38	3	3	4	5	5	6	6	7	8	8	9	9	10	10	10	11	11
39	3	3	4	5	5	6	6	7	8	8	9	9	10	10	11	11	12
40	3	3	4	5	5	6	7	7	8	8	9	9	10	10	11	11	12
41	3	3	4	5	6	6	7	7	8	8	9	10	10	11	11	12	12
42	3	4	4	5	6	6	7	7	8	9	9	10	10	11	11	12	12
43	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	12	13
44	3	4	4	5	6	6	7	8	8	9	9	10	11	11	12	17	13
45	3	4	4	5	6	7	7	8	8	9	10	10	11	11	12	12	13
46	3	4	4	5	6	7	7	8	9	9	10	10	11	11	12	13	13
47	3	4	5	5	6	7	7	8	9	9	10	10	11	12	12	13	13
48	3	4	5	5	6	7	7	8	9	9	10	11	11	12	12	13	14
49	3	4	5	5	6	7	8	8	9	10	10	11	11	12	13	13	14
50	3	4	5	5	6	7	8	8	9	10	10	11	11	12	13	13	14
ρ=0,01																	
26	3	4	4	5	5	6	6	7	7	8	8	8	9	9	10	10	10
27	3	4	4	5	6	6	6	7	7	8	8	9	9	9	10	10	11
28	3	4	4	5	6	6	7	7	8	8	8	9	9	10	10	10	11
29	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	11	11
30	3	4	5	5	6	6	7	7	8	8	9	9	10	10	10	11	11
31	3	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11	12
32	3	4	5	5	6	7	7	8	8	9	9	10	10	10	11	11	12
33	3	4	5	5	6	7	7	8	8	9	9	10	10	11	11	12	12
34	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	12	12
35	3	4	5	6	6	7	7	8	9	9	10	10	11	11	12	12	13
36	3	4	5	6	6	7	8	8	9	9	10	10	11	11	12	12	13
37	3	4	5	6	6	7	8	8	9	9	10	11	11	12	12	13	13
38	3	4	5	6	7	7	8	8	9	10	10	11	11	12	12	13	13
39	3	4	5	6	7	7	8	9	9	10	10	11	11	12	12	13	13
40	3	4	5	6	7	7	8	9	9	10	10	11	12	12	13	13	14
41	3	4	5	6	7	8	8	9	9	10	11	11	12	12	13	13	14
42	3	4	5	6	7	8	8	9	10	10	11	11	12	13	13	14	14
43	3	5	5	6	7	8	8	9	10	10	11	12	12	13	13	14	14
44	3	5	5	6	7	8	9	9	10	11	11	12	12	13	14	14	15
45	4	5	6	6	7	8	9	9	10	11	11	12	13	13	14	14	15
46	4	5	6	6	7	8	9	9	10	11	11	12	13	13	14	15	15
47	4	5	6	7	7	8	9	10	10	11	12	12	13	14	14	15	15
48	4	5	6	7	7	8	9	10	10	11	12	12	13	14	14	15	16
49	4	5	6	7	8	8	9	10	11	11	12	13	13	14	15	15	16
50	4	5	6	7	8	8	9	10	11	11	12	13	13	14	15	15	16

N	P 0,18 / Q 0,82	0,19 / 0,81	0,20 / 0,80	0,21 / 0,79	0,22 / 0,78	0,23 / 0,77	0,24 / 0,76	0,25 / 0,75	0,26 / 0,74	0,27 / 0,73	0,28 / 0,72	0,29 / 0,71	0,30 / 0,70	0,31 / 0,69	0,32 / 0,68	0,33 / 0,67	0,34 / 0,66
									p=0,05								
2	2	2	2	2	2	-	-	-	-	-	-	-	-	-	-	-	-
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
5	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
6	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5
7	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5
8	4	4	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6
9	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6
10	5	5	5	5	5	6	6	6	6	6	6	6	6	7	7	7	7
11	5	5	6	6	6	6	6	6	6	6	7	2	7	7	7	7	7
12	5	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8
13	6	6	6	6	6	7	7	7	7	7	7	8	8	8	8	8	8
14	6	6	6	7	7	7	7	7	7	8	8	8	8	8	8	9	9
15	6	6	7	7	7	7	7	8	8	8	8	8	9	9	9	9	9
16	7	7	7	7	7	8	8	8	8	8	9	9	9	9	9	9	10
17	7	7	7	7	8	8	8	8	8	9	9	9	9	9	10	10	10
18	7	7	8	8	8	8	8	9	9	9	9	9	10	10	10	10	10
19	7	8	8	8	8	9	9	9	9	9	10	10	10	10	10	11	11
20	8	8	8	8	9	9	9	9	10	10	10	10	10	11	11	11	11
21	8	8	8	9	9	9	10	10	10	10	10	11	11	11	11	12	12
22	8	8	9	9	9	10	10	10	10	10	11	11	11	11	12	12	12
23	8	9	9	9	9	10	10	10	11	11	11	11	12	12	12	12	13
24	9	9	9	9	10	10	10	11	11	11	11	12	12	12	13	13	13
25	9	9	9	10	10	10	11	11	11	12	12	12	12	13	13	13	13
									p=0,01								
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	3	3	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	-	-	-
5	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5
6	5	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6
7	5	5	5	6	6	6	6	6	6	6	6	6	6	6	6	6	6
8	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7
9	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7
10	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8
11	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9
12	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9
13	7	7	7	7	8	8	8	8	8	8	9	9	9	9	9	9	10
14	7	7	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10
15	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	10
16	8	8	8	8	9	9	9	9	9	10	10	10	10	10	11	11	11
17	8	8	9	9	9	9	9	10	10	10	10	10	11	11	11	11	11
18	8	9	9	9	9	10	10	10	10	10	11	11	11	11	12	12	12
19	9	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	12
20	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	13	13
21	9	10	10	10	10	11	11	11	11	12	12	12	12	13	13	13	13
22	10	10	10	10	11	11	11	11	12	12	12	13	13	13	13	14	14
23	10	10	10	11	11	11	12	12	12	12	13	13	13	13	14	14	14
24	10	10	11	11	11	12	12	12	13	13	13	13	14	14	14	14	15
25	10	11	11	11	12	12	12	13	13	13	13	14	14	14	15	15	15

N	P 0,18 / Q 0,82	0,19 / 0,81	0,20 / 0,80	0,21 / 0,79	0,22 / 0,78	0,23 / 0,77	0,24 / 0,76	0,25 / 0,75	0,26 / 0,74	0,27 / 0,73	0,28 / 0,72	0,29 / 0,71	0,30 / 0,70	0,31 / 0,69	0,32 / 0,68	0,33 / 0,67	0,34 / 0,66
									p=0,05								
26	9	9	10	10	10	11	11	11	12	12	12	12	13	13	13	14	14
27	9	10	10	10	11	11	11	12	12	12	12	13	13	13	14	14	14
28	10	10	10	11	11	11	12	12	12	13	13	13	13	14	14	14	15
29	10	10	10	11	11	12	12	12	13	13	13	14	14	14	14	15	15
30	10	10	11	11	11	12	12	13	13	13	14	14	14	15	15	15	16
31	10	11	11	11	12	12	12	13	13	14	14	14	15	15	15	16	16
32	10	11	11	12	12	12	13	13	14	14	14	15	15	15	16	16	16
33	11	11	12	12	12	13	13	13	14	14	15	15	15	16	16	16	17
34	11	11	12	12	13	13	13	14	14	15	15	15	16	16	16	17	17
35	11	12	12	12	13	13	14	14	14	15	15	16	16	16	17	17	18
36	11	12	12	13	13	14	14	14	15	15	16	16	16	17	17	18	18
37	12	12	13	13	13	14	14	15	15	16	16	16	17	17	18	18	18
38	12	12	13	13	14	14	15	15	15	16	16	17	17	18	18	18	19
39	12	13	13	14	14	14	15	15	16	16	17	17	17	18	18	19	19
40	12	13	13	14	14	15	15	16	16	17	17	17	18	18	19	19	20
41	13	13	14	14	15	15	15	16	16	17	17	18	18	19	19	20	20
42	13	13	14	14	15	15	16	16	17	17	18	18	19	19	19	20	20
43	13	14	14	15	15	16	16	17	17	18	18	18	19	19	20	20	21
44	13	14	14	15	15	16	16	17	17	18	18	19	19	20	20	21	21
45	13	14	15	15	16	16	17	17	18	18	19	19	20	20	21	21	22
46	14	14	15	15	16	16	17	17	18	18	19	20	20	21	21	22	22
47	14	15	15	16	16	17	17	18	18	19	19	20	20	21	21	22	22
48	14	15	15	16	16	17	18	18	19	19	20	20	21	21	22	22	23
49	14	15	16	16	17	17	18	18	19	19	20	21	21	22	22	23	23
50	14	15	16	16	17	17	18	18	19	19	20	21	21	22	22	23	23
									p=0,01								
26	11	11	11	12	12	12	13	13	13	14	14	14	14	15	15	15	16
27	11	11	12	12	12	13	13	13	14	14	14	15	15	15	15	16	16
28	11	12	12	12	13	13	13	14	14	14	15	15	15	16	16	16	17
29	11	12	12	13	13	13	14	14	14	15	15	15	16	16	16	17	17
30	12	12	12	13	13	14	14	14	15	15	15	16	16	16	17	17	17
31	12	12	13	13	14	14	14	15	15	15	16	16	16	17	17	17	18
32	12	13	13	13	14	14	15	15	15	16	16	16	17	17	18	18	18
33	12	13	13	14	14	15	15	15	16	16	16	17	17	18	18	18	19
34	13	13	14	14	14	15	15	16	16	16	17	17	17	18	18	19	19
35	13	13	14	14	15	15	16	16	16	17	17	18	18	18	19	19	20
36	13	14	14	15	15	15	16	16	17	17	18	18	18	19	19	20	20
37	13	14	14	15	15	16	16	17	17	18	18	18	19	19	20	20	20
38	14	14	15	15	16	16	17	17	17	18	18	19	19	20	20	20	21
39	14	14	15	15	16	16	17	17	18	18	19	19	20	20	20	21	21
40	14	15	15	16	16	17	17	18	18	19	19	20	20	20	21	21	22
41	14	15	16	16	17	17	18	18	18	19	19	20	20	21	21	22	22
42	15	15	16	16	17	17	18	18	19	19	20	20	21	21	22	22	23
43	15	16	16	17	17	18	18	19	19	20	20	21	21	22	22	23	23
44	15	16	16	17	17	18	18	19	19	20	21	21	21	22	22	23	23
45	15	16	17	17	18	18	19	19	20	20	21	21	22	22	23	23	24
46	16	16	17	17	18	19	19	20	20	21	21	22	22	23	23	24	24
47	16	17	17	18	18	19	19	20	21	21	22	22	23	23	24	24	25
48	16	17	17	18	19	19	20	20	21	21	22	22	23	24	24	25	25
49	16	17	18	18	19	19	20	21	21	22	22	23	23	24	24	25	26
50	16	17	18	18	19	19	20	21	21	22	22	23	23	24	24	25	26

N \ P/Q	0,35 / 0,65	0,36 / 0,64	0,37 / 0,63	0,38 / 0,62	0,39 / 0,61	0,40 / 0,60	0,41 / 0,59	0,42 / 0,58	0,43 / 0,57	0,44 / 0,56	0,45 / 0,55	0,46 / 0,54	0,47 / 0,53	0,48 / 0,52	0,49 / 0,51	0,50 / 0,50
								$p=0,05$								
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	4	4	4	4	4	4	4	4	4	4	4	4	-	-	-	-
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6
7	6	6	6	6	6	6	6	6	6	6	6	6	6	7	7	7
8	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7
9	7	7	7	7	7	7	7	7	7	7	8	8	8	8	8	8
10	7	7	7	7	7	8	8	8	8	8	8	8	8	8	8	9
11	8	8	8	8	8	8	8	8	8	9	9	9	9	9	9	9
12	8	8	8	8	8	9	9	9	9	9	9	9	9	10	10	10
13	8	9	9	9	9	9	9	9	10	10	10	10	10	10	10	10
14	9	9	9	9	9	10	10	10	10	10	10	11	11	11	11	11
15	9	10	10	10	10	10	10	10	11	11	11	11	11	11	12	12
16	10	10	10	10	10	11	11	11	11	11	11	12	12	12	12	12
17	10	10	11	11	11	11	11	12	12	12	12	12	12	13	13	13
18	11	11	11	11	11	12	12	12	12	12	13	13	13	13	13	13
19	11	11	12	12	12	12	12	13	13	13	13	13	13	14	14	14
20	12	12	12	12	12	13	13	13	13	13	14	14	14	14	14	15
21	12	12	12	13	13	13	13	14	14	14	14	14	15	15	15	15
22	12	13	13	13	13	14	14	14	14	15	15	15	15	15	16	16
23	13	13	13	14	14	14	14	15	15	15	15	16	16	16	16	16
24	13	14	14	14	14	15	15	15	15	16	16	16	16	17	17	17
25	14	14	14	15	15	15	15	16	16	16	16	17	17	17	17	18
								$p=0,01$								
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	5	5	5	5	5	-	-	-	-	-	-	-	-	-	-	-
6	6	6	6	6	6	6	6	6	6	6	6	6	-	-	-	-
7	6	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	8
9	8	8	8	8	8	8	8	8	8	8	8	9	9	9	9	9
10	8	8	8	8	9	9	9	9	9	9	9	9	9	9	9	10
11	9	9	9	9	9	9	9	9	10	10	10	10	10	10	10	10
12	9	9	9	10	10	10	10	10	10	10	10	10	11	11	11	11
13	10	10	10	10	10	10	10	11	11	11	11	11	11	11	11	12
14	10	10	10	11	11	11	11	11	11	11	12	12	12	12	12	12
15	11	11	11	11	11	11	12	12	12	12	12	12	12	13	13	13
16	11	11	11	12	12	12	12	12	12	13	13	13	14	14	14	14
17	12	12	12	12	12	13	13	13	13	13	13	14	14	14	14	14
18	12	12	13	13	13	13	13	13	14	14	14	14	14	14	15	15
19	13	13	13	13	13	14	14	14	14	14	15	15	15	15	15	15
20	13	13	14	14	14	14	14	15	15	15	15	15	16	16	16	16
21	14	14	14	14	14	15	15	15	15	16	16	16	16	16	17	17
22	14	14	15	15	15	15	15	16	16	16	16	17	17	17	17	17
23	15	15	15	15	15	16	16	16	16	17	17	17	17	18	18	18
24	15	15	15	16	16	16	16	17	17	17	17	18	18	18	18	19
25	15	16	16	16	17	17	17	17	18	18	18	18	19	19	19	19

N \ P/Q	0,35 / 0,65	0,36 / 0,64	0,37 / 0,63	0,38 / 0,62	0,39 / 0,61	0,40 / 0,60	0,41 / 0,59	0,42 / 0,58	0,43 / 0,57	0,44 / 0,56	0,45 / 0,55	0,46 / 0,54	0,47 / 0,53	0,48 / 0,52	0,49 / 0,51	0,50 / 0,50
								$p=0,05$								
26	14	14	15	15	15	16	16	16	16	17	17	17	17	18	18	18
27	15	15	15	15	16	16	16	17	17	17	17	18	18	18	18	19
28	15	15	16	16	16	16	17	17	17	13	18	18	18	19	19	19
29	15	16	16	16	17	17	17	18	18	18	18	19	19	19	20	20
30	16	16	17	17	17	17	18	18	18	19	19	19	20	20	20	20
31	16	17	17	17	18	18	18	19	19	19	20	20	20	20	21	21
32	17	17	17	18	18	18	19	19	19	20	20	20	21	21	21	22
33	17	17	18	18	19	19	19	20	20	20	21	21	21	22	22	22
34	18	18	18	19	19	19	20	20	20	21	21	21	22	22	22	23
35	18	18	19	19	19	20	20	21	21	21	22	22	22	23	23	23
36	18	19	19	20	20	20	21	21	21	22	22	22	23	23	24	24
37	19	19	20	20	20	21	21	22	22	22	23	23	23	24	24	24
38	19	20	20	20	21	21	22	22	22	23	23	24	24	24	25	25
39	20	20	20	21	21	22	22	22	23	23	24	24	24	25	25	26
40	20	20	21	21	22	22	23	23	23	24	24	25	25	25	26	26
41	20	21	21	22	22	23	23	23	24	24	25	25	26	26	26	27
42	21	21	22	22	23	23	23	24	24	25	25	26	26	26	27	27
43	21	22	22	23	23	24	24	24	25	25	26	26	27	27	27	26
44	22	22	23	23	24	24	24	25	25	26	26	27	27	28	28	28
45	22	23	23	24	24	24	25	25	26	26	27	27	28	28	29	29
46	22	23	23	24	24	25	25	26	26	27	27	28	28	29	29	30
47	23	23	24	24	25	25	26	26	27	27	28	28	29	29	30	30
48	23	24	24	25	25	26	26	27	27	28	28	29	29	30	30	31
49	24	24	25	25	26	26	27	27	28	28	29	29	30	30	31	31
50	24	24	25	25	26	26	27	27	28	28	29	29	30	30	31	32
								$p=0,01$								
26	16	16	16	17	17	17	18	18	18	18	19	19	19	19	20	20
27	16	17	17	17	18	18	18	18	19	19	19	19	20	20	20	20
28	17	17	17	18	18	18	19	19	19	19	20	20	20	21	21	21
29	17	18	18	18	19	19	19	19	20	20	20	21	21	21	21	22
30	18	18	18	19	19	19	20	20	20	21	21	21	21	22	22	22
31	18	19	19	19	19	20	20	20	21	21	21	22	22	22	23	23
32	19	19	19	20	20	20	21	21	21	22	22	22	22	23	23	24
33	19	19	20	20	20	21	21	21	22	22	22	23	23	23	24	24
34	20	20	20	21	21	21	22	22	22	23	23	23	24	24	24	25
35	20	20	21	21	21	22	22	23	23	23	24	24	24	25	25	25
36	20	21	21	22	22	22	23	23	23	24	24	25	25	25	26	26
37	21	21	22	22	22	23	23	24	24	24	25	25	25	26	26	27
38	21	22	22	23	23	23	24	24	24	25	25	26	26	26	27	27
39	22	22	23	23	23	24	24	25	25	25	26	26	27	27	27	28
40	22	23	23	23	24	24	25	25	26	26	26	27	27	28	28	28
41	23	23	23	24	24	25	25	26	26	26	27	27	28	28	28	29
42	23	23	24	24	25	25	26	26	27	27	27	28	28	29	29	29
43	23	24	24	25	25	26	26	27	27	28	28	28	29	29	30	30
44	24	24	25	25	26	26	27	27	28	28	28	29	29	30	30	31
45	24	25	25	26	26	27	27	28	28	29	29	29	30	30	31	31
46	25	25	26	26	27	27	28	28	29	29	30	30	30	31	31	32
47	25	26	26	27	27	28	28	29	29	30	30	31	31	31	32	32
48	26	26	27	27	28	28	29	29	30	30	31	31	32	32	33	33
49	26	27	27	28	28	29	29	30	30	31	31	32	32	33	33	34
50	26	27	27	28	28	29	29	30	30	31	31	32	32	33	33	34

Таблица XVI

Критические значения выборочного коэффициента корреляции рангов (по В.Ю. Урбаху, 1964)

Связь достоверна, если $r_{s\ эмп} \geq r_{s\ 0,05}$, и тем более достоверна, если $r_{s\ эмп} \geq r_{s\ 0,01}$

n	p 0,05	p 0,01	n	p 0,05	p 0,01	n	p 0,05	p 0,01
5	0,94	–	17	0,48	0,62	29	0,37	0,48
6	0,85	–	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,49	0,51	37	0,33	0,43
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40

Таблица XVII

Критические значения критерия F Фишера для уровней статистической значимости $p \leq 0,05$ и $p \leq 0,01$: df_1 - число степеней свободы в числителе, df_2 - число степеней свободы в знаменателе (по Snedecor G. B., 1956)

Влияние фактора или взаимодействия факторов достоверно, если $F_{эмп}$ равен или больше критического значения $F_{0,05}$ и тем более достоверно, если $F_{эмп} \geq F_{0,01}$.

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	11	12
	$p \leq 0,05$											
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
	$p \leq 0,01$											
1	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6082	6106
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,41	99,42
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,29	10,15	10,05	9,96	9,89
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71
11	9,65	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55

Таблица XVII. Продолжение

df_1 / df_2	1	2	3	4	5	6	7	8	9	10	11	12
						$p \leq 0{,}05$						
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38
18	4,41	3,55	3,16	2,93	2,77	3,66	2,58	2,51	2,46	2,41	2,37	2,34
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,34	2,31
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,26	2,23
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,24	2,20
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,20	2,16
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,16	2,13
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,15	2,12
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,14	2,10
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09
32	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,10	2,07
34	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,08	2,05
						$p \leq 0{,}01$						
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45
18	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,44	3,37
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30
20	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	3,30	3,23
21	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,40	3,31	3,24	3,17
22	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,25	3,17	3,09	3,03
25	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	3,05	2,99
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	3,02	2,96
27	7,68	5,49	4,60	4,11	3,79	3,56	3,39	3,26	3,14	3,06	2,98	2,93
28	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,95	2,90
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,08	3,00	2,92	2,87
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,90	2,84
32	7,50	5,34	4,46	3,97	3,66	3,42	3,25	3,12	3,01	2,94	2,86	2,80
34	7,44	5,29	4,42	3,93	3,61	3,38	3,21	3,08	2,97	2,89	2,82	2,76

Таблица XVII. Продолжение

df_1 / df_2	1	2	3	4	5	6	7	8	9	10	11	12
						$p \leq 0{,}05$						
36	4,11	3,26	2,86	2,63	2,48	2,36	2,28	2,21	2,15	2,10	2,06	2,03
38	4,10	3,25	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,05	2,02
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	2,02	1,99
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	2,01	1,98
46	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,14	2,09	2,04	2,00	1,97
48	4,04	3,19	2,80	2,56	2,41	2,30	2,21	2,14	2,08	2,03	1,99	1,96
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95
55	4,02	3,17	2,78	2,54	2,38	2,27	2,18	2,11	2,05	2,00	1,97	1,93
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92
65	3,99	3,14	2,75	2,51	2,36	2,24	2,15	2,08	2,02	1,98	1,94	1,90
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93	1,89
80	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99	1,95	1,91	1,88
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88	1,85
125	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,95	1,90	1,86	1,83
150	3,91	3,06	2,67	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85	1,82
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80
400	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,78
1000	3,85	3,00	2,61	2,38	2,22	2,10	2,02	1,95	1,89	1,84	1,80	1,76
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79	1,75
						$p \leq 0{,}01$						
36	7,39	5,25	4,38	3,89	3,58	3,35	3,18	3,04	2,94	2,86	2,78	2,72
38	7,35	5,21	4,34	3,86	3,54	3,32	3,15	3,02	2,91	2,82	2,75	2,69
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,88	2,80	2,73	2,66
42	7,27	5,15	4,29	3,80	3,49	3,26	3,10	2,96	2,86	2,77	2,70	2,64
44	7,24	5,12	4,26	3,78	3,46	3,24	3,07	2,94	2,84	2,75	2,68	2,62
46	7,21	5,10	4,24	3,76	3,44	3,22	3,05	2,92	2,82	2,73	2,66	2,60
48	7,19	5,08	4,22	3,74	3,42	3,20	3,04	2,90	2,80	2,71	2,64	2,58
50	7,17	5,06	4,20	3,72	3,41	3,18	3,02	2,88	2,78	2,70	2,62	2,56
55	7,12	5,01	4,16	3,68	3,37	3,15	2,98	2,85	2,75	2,66	2,59	2,53
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,56	2,50
65	7,04	4,95	4,10	3,62	3,31	3,09	2,93	2,79	2,70	2,61	2,54	2,47
70	7,01	4,92	4,08	3,60	3,29	3,07	2,91	2,77	2,67	2,59	2,51	2,45
80	6,96	4,88	4,04	3,56	3,25	3,04	2,87	2,74	2,64	2,55	2,48	2,41
100	6,90	4,82	3,98	3,51	3,20	2,99	2,82	2,69	2,59	2,51	2,43	2,36
125	6,84	4,78	3,94	3,47	3,17	2,95	2,79	2,65	2,56	2,47	2,40	2,33
150	6,81	4,75	3,91	3,44	3,14	2,92	2,76	2,62	2,53	2,44	2,37	2,30
200	6,76	4,71	3,88	3,41	3,11	2,90	2,73	2,60	2,50	2,41	2,34	2,28
400	6,70	4,66	3,83	3,36	3,06	2,85	2,69	2,55	2,46	2,37	2,29	2,23
1000	6,66	4,62	3,80	3,34	3,04	2,82	2,66	2,53	2,43	2,34	2,26	2,20
∞	6,64	4,60	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,24	2,18

df_1	14	16	20	24	30	40	50	75	100	200	500	∞
df_2	$p≤0,05$											
1	245	246	248	249	250	251	252	253	253	254	254	254
2	19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,48	19,49	19,49	19,50	19,50
3	8,71	8,69	8,66	8,64	8,62	8,60	8,58	8,57	8,56	8,54	8,54	8,53
4	5,87	5,84	5,80	5,77	5,74	5,71	5,70	5,68	5,66	5,65	5,64	5,63
5	4,64	4,60	4,56	4,53	4,50	4,46	5,44	4,42	4,40	4,38	4,37	4,36
6	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,72	3,71	3,69	3,68	3,67
7	3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,29	3,28	3,25	3,24	3,23
8	3,23	3,20	3,15	3,12	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,93
9	3,02	2,98	2,93	2,90	2,86	2,82	2,89	2,77	2,76	2,73	2,72	2,71
10	2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,61	2,59	2,56	2,55	2,54
11	2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,47	2,45	2,42	2,41	2,40
12	2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,36	2,35	2,32	2,31	2,30
13	2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,28	2,26	2,24	2,22	2,21
14	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13
15	2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,15	2,12	2,10	2,08	2,07
16	2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,09	2,07	2,04	2,02	2,01
	$p≤0,01$											
1	6142	6169	6208	6234	6261	6286	6302	6323	6334	6352	6361	6366
2	99,43	99,44	99,45	99,46	99,47	99,48	99,48	99,49	99,49	99,49	99,50	99,50
3	26,92	26,83	26,69	26,60	26,50	26,41	26,35	26,27	26,23	26,18	26,14	26,12
4	14,24	14,15	14,02	13,93	13,83	13,74	13,69	13,61	13,57	13,52	13,48	13,46
5	9,77	9,68	9,55	9,47	9,38	9,29	9,24	9,17	9,13	9,07	9,04	9,02
6	7,60	7,52	7,39	7,31	7,23	7,14	7,09	7,02	6,99	6,94	6,90	6,88
7	6,35	6,27	6,15	6,07	5,98	5,90	5,85	5,78	5,75	5,70	5,67	5,65
8	5,56	5,48	5,36	5,28	5,20	5,11	5,06	5,00	4,96	4,91	4,88	4,86
9	5,00	4,92	4,80	4,73	4,64	4,56	4,51	4,45	4,41	4,36	4,33	4,31
10	4,60	4,52	4,41	4,33	4,25	4,17	4,12	4,05	4,01	3,96	3,93	3,91
11	4,29	4,21	4,10	4,02	3,94	3,86	3,80	3,74	3,70	3,66	3,62	3,60
12	4,05	3,98	3,86	3,78	3,70	3,61	3,56	3,49	3,46	3,41	3,38	3,36
13	3,85	3,78	3,67	3,59	3,51	3,42	3,37	3,30	3,27	3,21	3,18	3,16
14	3,70	3,62	3,51	3,43	3,34	3,26	3,21	3,14	3,11	3,06	3,02	3,00
15	3,56	3,48	3,36	3,29	3,20	3,12	3,07	3,00	2,97	2,92	2,89	2,87
16	3,45	3,37	3,25	3,18	3,10	3,01	2,96	2,98	2,86	2,80	2,77	2,75

df_1	14	16	20	24	30	40	50	75	100	200	500	∞
df_2	$p≤0,05$											
17	2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,04	2,02	1,99	1,97	1,96
18	2,29	2,25	2,19	2,15	2,11	2,07	2,04	2,00	1,98	1,95	1,93	1,92
19	2,26	2,21	2,15	2,11	2,07	2,02	2,00	1,96	1,94	1,91	1,90	1,88
20	2,23	2,18	2,12	2,08	2,04	1,99	1,96	1,92	1,90	1,87	1,85	1,84
21	2,20	2,15	2,09	2,05	2,00	1,96	1,93	1,89	1,87	1,84	1,82	1,81
22	2,18	2,13	2,07	2,03	1,98	1,93	1,91	1,87	1,84	1,81	1,80	1,78
23	2,14	2,10	2,04	2,00	1,96	1,91	1,88	1,84	1,82	1,79	1,77	1,76
24	2,13	2,09	2,02	1798	1,94	1,89	1,86	1,82	1,80	1,76	1,74	1,73
25	2,11	2,06	2,00	1,96	1,92	1,87	1,84	1,80	1,77	1,74	1,72	1,71
26	2,10	2,05	1,99	1,95	1,90	1,85	1,82	1,78	1,76	1,72	1,70	1,69
27	2,08	2,03	1,97	1,93	1,88	1,84	1,80	1,76	1,74	1,71	1,68	1,67
28	2,06	2,02	1,96	1,91	1,87	1,81	1,78	1,75	1,72	1,69	1,67	1,65
29	2,05	2,00	1,94	1,90	1,85	1,80	1,77	1,73	1,71	1,68	1,65	1,64
30	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,72	1,69	1,66	1,64	1,62
32	2,02	1,97	1,91	1,86	1,82	1,76	1,74	1,69	1,67	1,64	1,61	1,59
34	2,00	1,95	1,89	1,84	1,80	1,74	1,71	1,67	1,64	1,61	1,59	1,57
	$p≤0,01$											
17	3,35	3,27	3,16	3,08	3,00	2,92	2,86	2,79	2,76	2,70	2,67	2,65
18	3,27	3,19	3,07	3,00	2,91	2,83	2,78	2,71	2,68	2,62	2,59	2,57
19	3,19	3,12	3,00	2,92	2,84	2,76	2,70	2,63	2,60	2,54	2,51	2,49
20	3,13	3,05	2,94	2,86	2,77	2,69	2,63	2,56	2,53	2,47	2,44	2,42
21	3,07	2,99	2,88	2,80	2,72	2,63	2,58	2,51	2,47	2,42	2,38	2,36
22	3,02	2,94	2,83	2,75	2,67	2,58	2,53	2,46	2,42	2,37	2,33	2,31
23	2,97	2,89	2,78	2,70	2,62	2,53	2,48	2,41	2,37	2,32	2,28	2,26
24	2,93	2,85	2,74	2,66	2,58	2,49	2,44	2,36	2,33	2,27	2,23	2,21
25	2,89	2,81	2,70	2,62	2,54	2,45	2,40	2,32	2,29	2,23	2,19	2,17
26	2,86	2,77	2,66	2,58	2,50	2,41	2,36	2,28	2,25	2,19	2,15	2,13
27	2,83	2,74	2,63	2,55	2,47	2,38	2,33	2,25	2,21	2,16	2,12	2,10
28	2,80	2,71	2,60	2,52	2,44	2,35	2,30	2,22	2,18	2,13	2,09	2,06
29	2,77	2,68	2,57	2,49	2,41	2,32	2,27	2,19	2,15	2,10	2,06	2,03
30	2,74	2,66	2,55	2,47	2,38	2,29	2,24	2,16	2,13	2,07	2,03	2,01
32	2,70	2,62	2,51	2,42	2,34	2,25	2,20	2,12	2,08	2,02	1,98	1,96
34	2,66	2,58	2,47	2,38	2,30	2,21	2,15	2,08	2,04	1,98	1,94	1,91

df_1	14	16	20	24	30	40	50	75	100	200	500	∞
df_2						$p \leq 0{,}05$						
36	1,98	1,93	1,87	1,82	1,78	1,72	1,69	1,65	1,62	1,59	1,56	1,55
38	1,96	1,92	1,85	1,80	1,76	1,71	1,67	1,63	1,60	1,57	1,54	1,53
40	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51
42	1,94	1,89	1,82	1,78	1,73	1,68	1,64	1,60	1,57	1,54	1,51	1,49
44	1,92	1,88	1,81	1,76	1,72	1,66	1,63	1,58	1,56	1,52	1,50	1,48
46	1,91	1,87	1,80	1,75	1,71	1,65	1,62	1,57	1,54	1,51	1,48	1,46
48	1,90	1,86	1,79	1,74	1,70	1,64	1,61	1,56	1,53	1,50	1,47	1,45
50	1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44
55	1,88	1,83	1,76	1,72	1,67	1,61	1,58	1,52	1,50	1,46	1,43	1,41
60	1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,50	1,48	1,44	1,41	1,39
65	1,85	1,80	1,73	1,68	1,63	1,57	1,54	1,49	1,46	1,42	1,39	1,37
70	1,84	1,79	1,72	1,67	1,62	1,56	1,53	1,47	1,45	1,40	1,37	1,35
80	1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,45	1,42	1,38	1,35	1,32
100	1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,42	1,39	1,34	1,30	1,28
125	1,77	1,72	1,65	1,60	1,55	1,49	1,45	1,39	1,36	1,31	1,27	1,25
150	1,76	1,71	1,64	1,59	1,54	1,47	1,44	1,37	1,34	1,29	1,25	1,22
200	1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,35	1,32	1,26	1,22	1,19
400	1,72	1,67	1,60	1,54	1,49	1,42	1,38	1,32	1,28	1,22	1,16	1,13
1000	1,70	1,65	1,58	1,53	1,47	1,41	1,36	1,30	1,26	1,19	1,13	1,08
∞	1,69	1,64	1,57	1,52	1,46	1,40	1,35	1,28	1,24	1,17	1,11	1,00
						$p \leq 0{,}01$						
36	2,62	2,54	2,43	2,35	2,26	2,17	2,12	2,04	2,00	1,94	1,90	1,87
38	2,59	2,51	2,40	2,32	2,22	2,14	2,08	2,00	1,97	1,90	1,86	1,84
40	2,56	2,49	2,37	2,29	2,20	2,11	2,05	1,97	1,94	1,88	1,84	1,81
42	2,54	2,46	2,35	2,26	2,17	2,08	2,02	1,94	1,91	1,85	1,80	1,78
44	2,52	2,44	2,32	2,24	2,15	2,06	2,00	1,92	1,88	1,82	1,78	1,75
46	2,50	2,42	2,30	2,22	2,13	2,04	1,98	1,90	1,86	1,80	1,76	1,72
48	2,48	2,40	2,28	2,20	2,11	2,02	1,96	1,88	1,84	1,78	1,73	1,70
50	2,46	2,39	2,26	2,18	2,10	2,00	1,94	1,86	1,82	1,76	1,71	1,68
55	2,43	2,35	2,23	2,15	2,06	1,96	1,90	1,82	1,78	1,71	1,66	1,64
60	2,40	2,32	2,20	2,12	2,03	1,93	1,87	1,79	1,74	1,68	1,63	1,60
65	2,37	2,30	2,18	2,09	2,00	1,90	1,84	1,76	1,71	1,64	1,60	1,56
70	2,35	2,28	2,15	2,07	1,98	1,88	1,82	1,74	1,69	1,62	1,56	1,53
80	2,32	2,24	2,11	2,03	1,94	1,84	1,78	1,70	1,65	1,57	1,52	1,49
100	2,26	2,19	2,06	1,98	1,89	1,79	1,73	1,64	1,59	1,51	1,46	1,43
125	2,23	2,15	2,03	1,94	1,85	1,75	1,68	1,59	1,54	1,46	1,40	1,37
150	2,20	2,12	2,00	1,91	1,83	1,72	1,66	1,56	1,51	1,43	1,37	1,33
200	2,17	2,09	1,97	1,88	1,79	1,69	1,62	1,53	1,48	1,39	1,33	1,28
400	2,12	2,04	1,92	1,84	1,74	1,64	1,57	1,47	1,42	1,32	1,24	1,19
1000	2,09	2,01	1,89	1,81	1,71	1,61	1,54	1,44	1,38	1,28	1,19	1,11
∞	2,07	1,99	1,87	1,79	1,69	1,59	1,52	1,41	1,36	1,25	1,15	1,00